NUCLEIC ACIDS, KITS, AND METHODS FOR THE DIAGNOSIS, PROGNOSIS AND TREATMENT OF GLAUCOMA AND RELATED DISORDERS

FIELD OF THE INVENTION

The present invention relates to the field of diagnostic and prognostic methods and kits, treatments, and compositions useful in understanding and identifying glaucoma, related intraocular pressure-disorders, and steroid sensitivity.

CROSS REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of U.S. Patent Application serial no. 08/938,669, filed September 26, 1997, specifically incorporated by reference herein, which is a continuation-in-part of U.S. Patent Application serial no. 08/791,154, filed January 28, 1997, also specifically incorporated by reference herein.

BACKGROUND OF THE INVENTION

A group of debilitating eye diseases, the "Glaucomas" represent the leading cause of preventable blindness in the United States and other developed nations. In general, glaucomas are characterized by the alteration of the trabecular meshwork (TM), which consists of specialized endothelial cells and their associated connective tissue. The TM endothelial cells line the path the aqueous humor of the eye filters through during the normal, physiological flux. The cells generate and regulate the TM by producing extracellular molecules, the composition of which is thought to directly control the aqueous fluid flow.

In Primary Open Angle Glaucoma ("POAG"), the most common form of glaucoma, an alteration in the TM leads to an obstruction of the normal ability of aqueous humor to leave its chamber surrounding the iris. However, the specific cells in the chamber between the iris and the cornea, in a region called the iridocorneal angle, remain "open" in that they continue to allow the egress of aqueous fluid (see, Vaughan, D. *et al.*, In: *General Ophthalmology*, Appleton & Lange, Norwalk, CT, pp. 213-230 (1992); and *Gray's Anatomy*, 37th Ed., Churchill Livingstone, London, pp. 1180-1190 (1989)). As a result of the alteration in the TM and the obstruction, an increased intraocular pressure ("IOP") can be observed. IOP can result in progressive visual loss and blindness if not treated appropriately and in a timely fashion.

Glaucomas are estimated to affect between 0.4% and 3.3% of all adults over 40 years old (Leske, M.C. *et al.*, *Amer. J. Epidemiol. 113*:1843-1846 (1986); Bengtsson, B., *Br. J. Ophthamol. 73*:483-487 (1989); Strong, N.P., *Ophthal. Physiol. Opt. 12*:3-7 (1992)). Moreover, the prevalence of the disease rises to over 6% of those 75 years or older (Strong, N.P., *Ophthal. Physiol. Opt. 12*:3-7 (1992)).

A link between steroid, corticosteroid, or glucocorticoid treatments and the increased IOP found in POAG disease has long been suspected. While only 5% of the normal population have high IOP increases in response to topical glucocorticoids, greater than 40-50% of similarly treated patients with POAG show a high IOP increase (16 mm Hg). In addition, an Open Angle Glaucoma may be induced by exposure to glucocorticoids. This observation has suggested that an increased or abnormal glucocorticoid response in trabecular cells of the TM may be involved in POAG (Zhan, G.L. *et al.*, *Exper. Eye Res. 54*:211-218 (1992); Yun, A.J. *et al., Invest. Ophthamol. Vis. Sci. 30*:2012-2022 (1989); Clark, A.F., *Exper. Eye Res. 55*:265 (1992); Klemetti, A., *Acta Ophthamol. 68*:29-33 (1990); Knepper, P.A., U.S. Patent No. 4,617,299).

The ability of glucocorticoids to induce a glaucoma-like condition has led to efforts to identify genes or gene products induced by the cells of the trabecular meshwork in response (Polansky, J.R. *et al.*, In: *Glaucoma Update IV*, Springer-Verlag, Berlin, pp. 20-29 (1991); Polansky J.R. and Weinrob, R.N., In: *Handbook of Experimetal Pharmacology, Vol. 69,* Springer-Verlag, Berlin, pp. 461-538 (1984)). Initial efforts using short-term exposure to dexamethasone revealed only changes in specific protein synthesis. Extended exposure to relatively high levels of dexamethasone was, however, found to induce the expression of related 66 kD and 55 kD proteins that could be visualized by gel electrophoresis (Polansky, J.R. *et al.*, In: *Glaucoma Update IV*, Springer-Verlag, Berlin, pp. 20-29 (1991)). The induction kinetics of these proteins as well as their dose response characteristics were similar to the kinetics that were required for steroid-induced IOP elevation in human subjects (Polansky, J.R. *et al.*, In: *Glaucoma Update IV*, Springer-Verlag, Berlin, pp. 20-29 (1991)). Problems of aggregation and apparent instability or loss of protein in the purification process were obstacles in obtaining a direct protein sequence.

Nguyen *et al.,* U.S. Patent Application No: 08/649,432, filed May 17, 1996, now U.S. Patent No. 5,789,169, the entire disclosure of which is hereby incorporated by reference as if set forth at length herein, disclosed a novel protein sequence (the TIGR, trabecular meshwork inducible glucocorticoid response protein) highly induced by glucocorticoids in the endothelial lining cells of the human trabecular meshwork. Nguyen *et al.* also disclosed the cDNA sequence for that protein, the protein itself, molecules that bind to it, and nucleic acid molecules that encode

it, and provided improved methods and reagents for diagnosing glaucoma and related disorders, as well as for diagnosing other diseases or conditions, such as cardiovascular, immunological, or other diseases or conditions that affect the expression or activity of the protein.

Because increased IOP is a readily measurable characteristic of glaucoma, the diagnosis of the disease is largely screened for by measuring intraocular pressure (tonometry) (Strong, N.P., *Ophthal. Physiol. Opt. 12*:3-7 (1992), Greve, M. *et al., Can. J. Ophthamol. 28*:201-206 (1993)). Unfortunately, because glaucomatous and normal pressure ranges overlap, such methods are of limited value unless multiple readings are obtained (Hitchings, R.A., *Br. J. Ophthamol. 77*:326 (1993); Tuck, M.W. *et al., Ophthal. Physiol. Opt. 13*:227-232 (1993); Vaughan, D. *et al.,* In: *General Ophthamology*, Appleton & Lange, Norwalk, CT, pp. 213-230 (1992); Vernon, S.A., *Eye 7*:134-137 (1993)). Patients may also have a differential sensitivity to optic nerve damage at a given IOP. For these reasons, additional methods, such as direct examination of the optic disk and determination of the extent of a patient's visual field loss are often conducted to improve the accuracy of diagnosis (Greve, M. *et al., Can. J. Ophthamol. 28*:201- 206 (1993)). Moreover, these techniques are of limited prognostic value. In some aspects, the present invention fulfills the need for improved diagnostic and prognostic methods.

The elevation of intraocular pressure (IOP) due to topical corticosteroids (and other routes of administration) is an important clinical problem that limits the clinical use of these effective anti-inflammatory agents. If not observed in sufficient time, the IOP elevation (especially in certain individuals who show the high end of steroid-induced IOP elevations) can result in optic nerve damage and permanent visual field loss, termed "steroid glaucoma." Even patients taking inhaled, nasal, rectal, and facial steroids may be at risk. The present invention, in part, provides improved diagnostic agents, prognostic agents, therapeutic agents and methods that address this clinical problem.

SUMMARY OF THE INVENTION

The invention relates to nucleic acids, genes, proteins and cells that can be used in the treatment, diagnosis, prognosis, and identification of glaucoma, IOP-related disorders, or steroid sensitivity. The invention encompasses numerous agents, compositions, and methods, some of which are described by the objects and aspects of the invention detailed below. Others can be devised from the entire contents of this disclosure, and from the detailed description that follows.

In one aspect, the invention relates to nucleic acids comprising non-coding regions or promoter regions associated with the TIGR (trabecular meshwork inducible glucocorticoid response) gene of mammals. These nucleic acids can be used in identifying polymorphisms in the

genomes of mammals and humans that predict steroid sensitivity or a susceptibility to glaucomas or diseases related to alterations in IOP. A number of diagnostic or prognostic methods and kits can be designed from these nucleic acids.

In one embodiment, the nucleic acids can be used to identify or detect a single base polymorphism in a genome. In other embodiments, two or more single base polymorphisms or multiple base polymorphisms can be identified or detected. The detection of a known polymorphism can be the basis for diagnostic and prognostic methods and kits of the invention. Various methods of detecting nucleic acids can be used in these methods and with the kits, including, but not limited to, solution hybridization, hybridization to microarrays containing immobilized nucleic acids or other immobilized nucleic acids, amplification-based methods such as PCR and the like, and an appropriate biosensor apparatus comprising a nucleic acid or nucleic acid binding reagent.

In another aspect, the invention relates to specific sequences and variants or mutants from the promoter or 5' regulatory region of the human TIGR gene and nucleic acids incorporating these sequences, variants or mutants. The nucleic acids can be incorporated into the methods and kits of the invention, or used in expression systems, vectors, and cells to produce a protein or polypeptide of interest, or used in methods to identify or detect regulatory proteins or proteins that specifically bind to promoter or regulatory regions of the TIGR gene. While many of the examples below detail work from human tissue, other animals may be used as a source of the sequences. In one embodiment of this aspect of the invention, for example, nucleic acids having the disclosed TIGRmt11 sequence variant, represented by the change at nucleotide 5113 in SEQ ID NO: 1, 3, or 34 from T to C, or the change in nucleotide 5117 in SEQ ID NO: 2 from T to C. The presence of sequence variant mt11 is linked to the high IOP response to steroid treatments and a nucleic acid incorporating the singe base substitution can be used to identify and determine individuals at risk for developing glaucoma from undergoing a steroid treatment therapy, or a progression from an ocular hypertensive state, or those with a steroid sensitivity. And, because of the link between high IOP responses to steroids and the later development of glaucoma, nucleic acids having the TIGRmt11 sequence variant may also be used to identify the risk of developing glaucomas, such as POAG. The identification of changes in IOP can be done by any known means, however, the "Armaly" criteria is preferred (*see* Armaly, M.F., *Arch. Ophthalmol.* 70:492 (1963); Armaly, M.F., *Arch Ophtalmol.* 75:32-35 (1966); Kitazawa, Y. *et al., Arch. Ophthalmol.* 99:819-823 (1981); Lewis, J.M. *et al., Amer. J. Ophthalmol.* 106:607-612 (1988); Becker, B. *et al. Amer. J.Ophthalmol.* 57:543 (1967), all of which are specifically incorporated herein by reference in their entireties).

An object of the invention is to provide a method for diagnosing glaucoma in a patient which comprises the steps: (A) incubating under conditions permitting nucleic acid hybridization: a marker nucleic acid molecule, said marker nucleic acid molecule comprising a nucleotide sequence of a polynucleotide that specifically hybridizes to a polynucleotide that is linked to a TIGR promoter, and a complementary nucleic acid molecule obtained from a cell or a bodily fluid of said patient, wherein nucleic acid hybridization between said marker nucleic acid molecule, and said complementary nucleic acid molecule obtained from said patient permits the detection of a polymorphism whose presence is predictive of a mutation affecting TIGR response in said patient; (B) permitting hybridization between said marker nucleic acid molecule and said complementary nucleic acid molecule obtained from said patient; and (C) detecting the presence of said polymorphism, wherein the detection of the polymorphism is diagnostic of glaucoma.

Another object of the invention is to provide a method for prognosing glaucoma in a patient which comprises the steps: (A) incubating under conditions permitting nucleic acid hybridization: a marker nucleic acid molecule, said marker nucleic acid molecule comprising a nucleotide sequence of a polynucleotide that specifically hybridizes to a polynucleotide that is linked to a TIGR promoter, and a complementary nucleic acid molecule obtained from a cell or a bodily fluid of said patient, wherein nucleic acid hybridization between said marker nucleic acid molecule, and said complementary nucleic acid molecule obtained from said patient permits the detection of a polymorphism whose presence is predictive of a mutation affecting TIGR response in said patient; (B) permitting hybridization between said marker nucleic acid molecule and said complementary nucleic acid molecule obtained from said patient; and (C) detecting the presence of said polymorphism, wherein the detection of the polymorphism is prognostic of glaucoma.

Another object of the invention is to provide marker nucleic acid molecules capable of specifically detecting *TIGRmt1, TIGRmt2, TIGRmt3, TIGRmt4, TIGRmt5, TIGRmt11* and *TIGRsv1* .

Another object of the invention is to provide a method for diagnosing steroid sensitivity in a patient which comprises the steps: (A) incubating under conditions permitting nucleic acid hybridization: a marker nucleic acid molecule, the marker nucleic acid molecule comprising a nucleotide sequence of a polynucleotide that is linked to a TIGR promoter, and a complementary nucleic acid molecule obtained from a cell or a bodily fluid of the patient, wherein nucleic acid hybridization between the marker nucleic acid molecule, and the complementary nucleic acid molecule obtained from the patient permits the detection of a polymorphism whose presence is predictive of a mutation affecting TIGR response in the patient; (B) permitting hybridization between said TIGR-encoding marker nucleic acid molecule and the complementary nucleic acid molecule obtained from the patient; and (C) detecting the presence of the polymorphism, wherein the detection of the polymorphism is diagnostic of steroid sensitivity.

Further objects of the invention provide a nucleic acid molecule that comprises the sequence of SEQ ID NO: 1 or 34, recombinant DNA molecules containing a polynucleotide that specifically hybridizes to SEQ ID NO: 1 or 34 and substantially purified molecules that specifically bind to a nucleic acid molecule that comprises the sequence of SEQ ID NO: 1 or 34.

Further objects of the invention provide a nucleic acid molecule that comprises the sequence of SEQ ID NO: 3, recombinant DNA molecules containing a polynucleotide that specifically hybridizes to SEQ ID NO: 3 and substantially purified molecules that specifically bind to a nucleic acid molecule that comprises the sequence of SEQ ID NO: 3.

Additional objects of the invention provide a nucleic acid molecule that comprises the sequence of SEQ ID NO: 4, recombinant DNA molecules containing a polynucleotide that specifically hybridizes to SEQ ID NO: 4 and substantially purified molecules that specifically bind to a nucleic acid molecule that comprises the sequence of SEQ ID NO: 4.

Additional objects of the invention provide a nucleic acid molecule that comprises the sequence of SEQ ID NO: 5, recombinant DNA molecules containing a polynucleotide that specifically hybridizes to SEQ ID NO: 5 and substantially purified molecules that specifically bind to a nucleic acid molecule that comprises the sequence of SEQ ID NO: 5.

An additional object of the present invention is to provide a method of treating glaucoma which comprises administering to a glaucomatous patient an effective amount of an agent that inhibits the synthesis of a TIGR protein.

Indeed, the molecules of the present invention may be used to diagnose diseases or conditions which are characterized by alterations in the expression of extracellular proteins.

BRIEF DESCRIPTION OF THE FIGURES:

Figures 1A, 1B, 1C, 1D and 1E provide the nucleic acid sequence of a TIGR promoter region (SEQ ID NO: 1) from an individual without glaucoma.

Figures 2A, 2B, 2C and 2D provide the location and sequence changes highlighted in bold associated with glaucoma mutants *TIGRmt1*, *TIGRmt2*, *TIGRmt3*, *TIGRmt4*, *TIGRmt5*, *TIGRmt11*, and *TIGRsv1* (SEQ ID NO: 2).

Figures 3A, 3B, 3C, 3D, 3E, 3F, and 3G provide nucleic acid sequences of a TIGR promoter, and TIGR exons, TIGR introns and TIGR downstream sequences (SEQ ID NO: 3, SEQ ID NO: 4, and SEQ ID NO: 5).

Figure 4 provides a diagrammatic representation of the location of primers on the TIGR gene promoter for Single Strand Conformational Polymorphism (SSCP) analysis.

Figure 5 provides a diagrammatic representation of the TIGR exons and the arrangement of SSCP primers.

Figure 6 provides a homology analysis of TIGR homology with olfactomedin and olfactomedin-related proteins.

Figure 7 shows the nucleotide sequence of TIGR (SEQ ID NO: 26).

Figure 8 shows the amino acid sequence of TIGR (SEQ ID NO: 32).

DETAILED DESCRIPTION OF THE INVENTION

I. Agents of the Invention

As used herein, the term "glaucoma" has its art recognized meaning, and includes both primary glaucomas, secondary glaucomas, juvenile glaucomas, congenital glaucomas, and familial glaucomas, including, without limitation, pigmentary glaucoma, high tension glaucoma and low tension glaucoma and their related diseases. The methods of the present invention are particularly relevant to the diagnosis of POAG, OAG, juvenile glaucoma, and inherited glaucomas. The methods of the present invention are also particularly relevant to the prognosis of POAG, OAG, juvenile glaucoma, and inherited glaucomas. A disease or condition is said to be related to glaucoma if it possesses or exhibits a symptom of glaucoma, for example, an increased intra-ocular pressure resulting from aqueous outflow resistance (see, Vaughan, D. *et al.*, In: *General Ophthamology*, Appleton & Lange, Norwalk, CT, pp. 213-230 (1992)). The preferred agents of the present invention are discussed in detail below.

The agents of the present invention are capable of being used to diagnose the presence or severity of glaucoma and its related diseases in a patient suffering from glaucoma (a "glaucomatous patient"). The agents of the present invention are also capable of being used to prognose the presence or severity of glaucoma and its related diseases in a person not yet suffering from any clinical manifestations of glaucoma. Such agents may be either naturally occurring or non-naturally occurring. As used herein, a naturally occurring molecule may be "substantially purified," if desired, such that one or more molecules that is or may be present in a naturally occurring preparation containing that molecule will have been removed or will be present at a lower concentration than that at which it would normally be found.

The agents of the present invention will preferably be "biologically active" with respect to either a structural attribute, such as the capacity of a nucleic acid to hybridize to another nucleic acid molecule, or the ability of a protein to be bound by antibody (or to compete with another molecule for such binding). Alternatively, such an attribute may be catalytic, and thus involve the capacity of the agent to mediate a chemical reaction or response.

As used herein, the term "TIGR protein" refers to a protein having the amino acid sequence of SEQ ID NO: 32. As used herein, the agents of the present invention comprise nucleic acid molecules, proteins, and organic molecules.

As indicated above, the trabecular meshwork has been proposed to play an important role in the normal flow of the aqueous, and has been presumed to be the major site of outflow resistance in glaucomatous eyes. Human trabecular meshwork (HTM) cells are endothelial like cells which line the outflow channels by which aqueous humor exits the eye; altered synthetic function of the cells may be involved in the pathogenesis of steroid glaucoma and other types of glaucoma. Sustained steroid treatment of these cells are interesting because it showed that a major difference was observed when compared to 1-2 day glucocorticoid (GC) exposure. This difference appears relevant to the clinical onset of steroid glaucoma (1-6 weeks).

Although trabecular meshwork cells had been found to induce specific proteins in response to glucocorticoids (see, Polansky, J.R., In: "*Basic Aspects of Glaucoma Research III*", *Schattauer, New York* 307-318 (1993)), efforts to purify the expressed protein were encumbered by insolubility and other problems. Nguyen, T.D. *et al.*, (In: "*Basic Aspects of Glaucoma Research III*", *Schattauer, New York*, 331-343 (1993), herein incorporated by reference) used a molecular cloning approach to isolate a highly induced mRNA species from glucocorticoid-induced human trabecular cells. The mRNA exhibited a time course of induction that was similar to the glucocorticoid-induced proteins. The clone was designated "II.2" (ATCC No: 97994, American Type Culture Collection, Manassas, VA).

Nguyen *et al.*, U.S. Patent Application No: 08/649,432 filed May 17, 1996, Patent No. 5789169, isolated a II.2 clone which encoded a novel secretory protein that is induced in cells of the trabecular meshwork upon exposure to glucocorticoids. It has been proposed that this protein may become deposited in the extracellular spaces of the trabecular meshwork and bind to the surface of the endothelial cells that line the trabecular meshwork, thus causing a decrease in aqueous flow. Quantitative dot blot analysis and PCR evaluations have shown that the mRNA exhibits a progressive induction with time whereas other known GC-inductions from other systems and found in HTM cells (metallothionein, alpha-1 acid glycoprotein and alpha-1 antichymotrypsin) reached maximum level at one day or earlier. Of particular interest, the induction level of this clone was very high (4-6% total cellular mRNA) with control levels undetectable without PCR method. Based on studies of ^{35}S methionine cell labeling, the clone has the characteristics recently discovered for the major GC-induced extracellular glycoprotein in these cells, which is a sialenated, N-glycosylated molecule with a putative inositol phosphate anchor. The induction of mRNA approached 4% of the total cellular mRNA. The mRNA increased progressively over 10 days of dexamethasone treatment. The II.2 clone is 2.0 Kb whereas the Northern blotting shows a band of 2.5 Kb. Although not including a poly A tail, the 3' end of the clone contains two consensus polyadenylation signals.

A genomic clone was isolated and designated P_1TIGR clone (ATCC No: 97570, American Type Culture Collection, Rockville, Maryland). In-situ hybridization using the P_1TIGR

clone shows a TIGR gene and/or a sequence or sequences that specifically hybridize to the TIGR gene located at chromosome 1, q21-27, and more preferably to the TIGR gene located at chromosome 1, q22-26, and most preferably to the TIGR gene located at chromosome 1, q24. Clone P_1TIGR comprises human genomic sequences that specifically hybridize to the TIGR gene cloned into the *Bam*HI site of vector pCYPAC (Ioannou *et al., Nature Genetics, 6:*84-89 (1994) herein incorporated by reference).

As used herein, the term "TIGR gene" refers to the region of DNA involved in producing a TIGR protein; it includes, without limitation, regions preceeding and following the coding region as well as intervening sequences between individual coding regions.

As used herein, the term "TIGR exon" refers to any interrupted region of the TIGR gene that serves as a template for a mature TIGR mRNA molecule. As used herein, the term "TIGR intron" refers to a region of the TIGR gene which is non-coding and serves as a template for a TIGR mRNA molecule.

Localization studies using a Stanford G3 radiation hybrid panel mapped the TIGR gene near the D1S2536 marker with a LOD score of 6.0 (Richard *et al., American Journal of Human Genetics 52.5:* 915-921 (1993), herein incorporated by reference); Frazer *et al., Genomics 14.3*: 574-578 (1992), herein incorporated by reference; Research Genetics, Huntsville, Alabama). Other markers in this region include: D1S210; D1S1552; D1S2536; D1S2790; SHGC-12820; and D1S2558.

Sequences located upstream of the TIGR coding region are isolated and sequenced in a non-glaucomic individual. The upstream sequence is set forth in SEQ ID. No. 1 and 34. Sequence comparisons of the upstream region of a non-glaucoma individual and individuals with glaucoma identify a number of mutations in individuals with glaucoma. Some of these mutations are illustrated in Figure 2, the sequence of which can be used to identify the exact changes in the human genomic sequences from the upstream region of the TIGR gene disclosed here (SEQ ID NO: 1, 2, 3, and 34). SEQ ID NO: 3 includes the regions through the start of transcription and the start of translation, as evident from a sequence comparison to the figures. SEQ ID NO: 34 ends before the transcription start site, again as evident from a sequence comparison with the figures. Six mutations are specifically disclosed here. *TIGRmt1* is the result of a replacement of a cytosine with a guanine at position 4337 (SEQ ID NO: 1, SEQ ID NO: 2, and SEQ ID NO: 3). *TIGRmt2* is the result of a replacement of a cytosine with a thymine at position 4950 (SEQ ID NO: 1, SEQ ID NO: 2, and SEQ ID NO: 3). *TIGRmt3* is the result of an addition in the following order of a guanine, a thymine, a guanine, and a thymine (GTGT) at position 4998 (SEQ ID NO: 1, SEQ ID NO: 2, and SEQ ID NO: 3). *TIGRmt4* is the result of a replacement of an adenine with a guanine at position 4256 (SEQ ID NO: 1, SEQ ID NO: 2, and SEQ ID NO: 3). *TIGRmt5* is the result of a replacement of a guanine with an adenine at position 4262 (SEQ ID

NO: 1, SEQ ID NO: 2 and SEQ ID NO: 3). *TIGRmt11* (not pictured in Figure 2) is the result of a replacement of a thymine with a cytosine at position 5113 (SEQ ID NO: 1, 3, or 34) and at the equivalent position in SEQ ID NO: 2, at nucleotide 5117. One or more of *TIGRmt1*, *TIGRmt2*, *TIGRmt3*, *TIGRmt4*, *TIGRmt5*, and *TIGTmt11* can be homozygous or heterozygous.

Sequence comparisons of the upstream region of a non-glaucoma individual and individuals with glaucoma identify at least one sequence variation in individuals with glaucoma. One such sequence variant is illustrated in Figure 2. *TIGRsv1* is the result of a replacement of an adenine with a guanine at position 4406 (SEQ ID NO: 1, SEQ ID NO: 2 and SEQ ID NO: 3). Also, the presence of *TIGRmt11* is associated with steroid sensitivity or an increased susceptibility to developing glaucoma or IOP-related disorders during steroid or corticosteroid treatment.

Molecules comprising sequences upstream of the TIGR coding region provide useful markers for polymorphic studies. Such molecules include primers suitable for single strand conformational polymorphic studies, examples of which are as follows: forward primer “Sk-1a”: 5’-TGA GGC TTC CTC TGG AAA C-3’ (SEQ ID NO: 6); reverse primer “ca2”: 5’-TGA AAT CAG CAC ACC AGT AG-3’ (SEQ ID NO: 7); forward primer “CA2”: 5’-GCA CCC ATA CCC CAA TAA TAG-3’ (SEQ ID NO: 8); reverse primer “Pr+1”: 5’-AGA GTT CCC CAG ATT TCA CC-3’ (SEQ ID NO: 9); forward primer “Pr-1”: 5’-ATC TGG GGA ACT CTT CTC AG-3’ (SEQ ID NO: 10); reverse primer ”Pr+2(4A2)”: 5’-TAC AGT TGT TGC AGA TAC G-3’ (SEQ ID NO: 11); forward primer “Pr-2(4A)”: 5’-ACA ACG TAT CTG CAA CAA CTG-3’ (SEQ ID NO: 12); reverse primer “Pr+3(4A)”: 5’-TCA GGC TTA ACT GCA GAA CC-3’ (SEQ ID NO: 13); forward primer “Pr-3(4A)”: 5’-TTG GTT CTG CAG TTA AGC C-3’ (SEQ ID NO: 14); reverse primer “Pr+2(4A1)”: 5’-AGC AGC ACA AGG GCA ATC C-3’ (SEQ ID NO: 15); reverse primer “Pr+1(4A)”: 5’-ACA GGG CTA TAT TGT GGG-3’ (SEQ ID NO: 16).

In addition, molecules comprising sequences within TIGR exons provide useful markers for polymorphic studies. Such molecules include primers suitable for single strand conformational polymorphic studies, examples of which are as follows: forward primer “KS1X”: 5’-CCT GAG ATG CCA GCT GTC C-3’ (SEQ ID NO: 17); reverse primer “SK1XX”: 5’-CTG AAG CAT TAG AAG CCA AC-3’ (SEQ ID NO: 18); forward primer “KS2a1”: 5’-ACC TTG GAC CAG GCT GCC AG-3’ (SEQ ID NO: 19); reverse primer “SK3” 5’-AGG TTT GTT CGA GTT CCA G-3’ (SEQ ID NO: 20); forward primer “KS4”: 5’-ACA ATT ACT GGC AAG TAT GG-3’ (SEQ ID NO: 21); reverse primer “SK6A”: 5’-CCT TCT CAG CCT TGC TAC C-3’ (SEQ ID NO: 22); forward primer “KS5”: 5’-ACA CCT CAG CAG ATG CTA CC-3’ (SEQ ID NO: 23); reverse primer “SK8”: 5’-ATG GAT GAC TGA CAT GGC C-3’ (SEQ ID NO: 24); forward primer “KS6”: 5’-AAG GAT GAA CAT GGT CAC C-3’ (SEQ ID NO: 25).

The locations of primers: Sk-1a, ca2, CA2, Pr+1, Pr-1, Pr+2(4A2), Pr-2(4A), Pr+3(4A), Pr-3(4A), Pr-3(4A), Pr+2(4A1), and Pr+1(4A) are diagrammatically set forth in Figure 4. The

location of primers: KS1X, SK1XX, Ks2a1, SK3, KS4, SK6A, KS5, SK8, and KS6 are diagramatically set forth in Figure 5.

The primary structure of the TIGR coding region initiates from an ATG initiation site (SEQ ID NO:3, residues 5337-5339) and includes a 20 amino acid consensus signal sequence a second ATG (SEQ ID NO: 3, residues 5379-5381), indicating that the protein is a secretory protein. The nucleotide sequence for the TIGR coding region is depicted in Figure 7 (SEQ ID NO: 26). The protein contains an N-linked glycosylation site located in the most hydrophilic region of the molecule. The amino terminal portion of the protein is highly polarized and adopts alpha helical structure as shown by its hydropathy profile and the Garnier-Robison structure analysis. In contrast, the protein contains a 25 amino acid hydrophobic region near its carboxy terminus. This region may comprise a glucocorticoid-induced protein (GIP) anchoring sequence. The amino acid sequence of TIGR is depicted in Figure 8 (SEQ ID NO: 32).

Study of cyclohexamide treatment in the absence and presence of GC suggest that the induction of TIGR may involve factors in addition to the GC receptor. The TIGR gene may be involved in the cellular stress response since it is also induced by stimulants such as H_2O_2, 12-O-tetradecanolyphorbol-13-acetate (TPA), and high glucose; this fact may relate to glaucoma pathogenesis and treatment.

A preferred class of agents comprises TIGR nucleic acid molecules ("TIGR molecules") or fragments thereof. Such molecules may be either DNA or RNA. A second preferred class of agents ("TIGR molecules") comprises the TIGR protein, its peptide fragments, fusion proteins, and analogs.

TIGR nucleic acid molecules or fragments thereof are capable of specifically hybridizing to other nucleic acid molecules under certain circumstances. As used herein, two nucleic acid molecules are said to be capable of specifically hybridizing to one another if the two molecules are capable of forming an anti-parallel, double-stranded nucleic acid structure. A nucleic acid molecule is said to be the "complement" of another nucleic acid molecule if the molecules exhibit complete complementarity. As used herein, molecules are said to exhibit "complete complementarity" when every nucleotide of one of the molecules is complementary to a nucleotide of the other. Two molecules are said to be "minimally complementary" if they can hybridize to one another with sufficient stability to permit them to remain annealed to one another under at least conventional "low-stringency" conditions. Similarly, the molecules are said to be "complementary" if they can hybridize to one another with sufficient stability to permit them to remain annealed to one another under conventional "high-stringency" conditions. Conventional stringency conditions are described by Sambrook et al., In: Molecular Cloning, A Laboratory Manual, 2nd Edition, Cold Spring Harbor Press, Cold Spring Harbor, New York (1989)), and by Haymes et al., In: Nucleic Acid Hybridization, A Practical Approach, IRL Press, Washington, DC

(1985), the entirety of which is herein incorporated by reference. Departures from complete complementarity are therefore permissible, as long as such departures do not completely preclude the capacity of the molecules to form a double-stranded structure. In order for an nucleic acid molecule to serve as a primer or probe it need only be sufficiently complementary in sequence to be able to form a stable double-stranded structure under the particular solvent and salt concentrations employed.

Appropriate stringency conditions which promote DNA hybridization, for example, 6.0 x sodium chloride/sodium citrate (SSC) at about 45°C, followed by a wash of 2.0 x SSC at 50°C, are known to those skilled in the art or can be found in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. For example, the salt concentration in the wash step can be selected from a low stringency of about 2.0 x SSC at 50°C to a high stringency of about 0.2 x SSC at 50°C. In addition, the temperature in the wash step can be increased from low stringency conditions at room temperature, about 22°C, to high stringency conditions at about 65°C. Both temperature and salt may be varied, or either the temperature or the salt concentration may be held constant while the other variable is changed.

In a preferred embodiment, a nucleic acid of the present invention will specifically hybridize to one or more of the nucleic acid molecules set forth in SEQ ID NO: 1-5 or 34, or complements thereof, or fragments of about 20 to about 200 bases of either, under moderately stringent conditions, for example at about 2.0 x SSC and about 65°C. In a particularly preferred embodiment, a nucleic acid of the present invention will specifically hybridize to one or more of the nucleic acid molecules set forth in SEQ ID NO: 1-5 or 34, or complements or fragments of either under high stringency conditions.

In one aspect of the present invention, a preferred marker nucleic acid molecule of the present invention has the nucleic acid sequence set forth in SEQ ID NO: 6-25 or 33, or complements thereof or fragments of either. In another aspect of the present invention, a preferred marker nucleic acid molecule of the present invention shares between about 80% to about 100% or about 90% to about 100% sequence identity with the nucleic acid sequence set forth in SEQ ID NO: 6 –25 or 33, or complement thereof or fragments of either. In a further aspect of the present invention, a preferred marker nucleic acid molecule of the present invention shares between about 95% to about 100% sequence identity with the sequence set forth in SEQ ID NO: 6-25 or 33, or complement thereof or fragments of either. In a more preferred aspect of the present invention, a preferred marker nucleic acid molecule of the present invention shares between 98% and about 100% sequence identity with the nucleic acid sequence set forth in SEQ ID NO:6-25 or 33, or complement thereof or fragments of either.

Regulatory Regions and Agents that Bind to the Regions or Agents that Alter the Binding of a Molecule that Binds to the Regions

Sequence comparisons of the upstream region identify a number of DNA motifs (*cis* elements) or regulatory regions. These DNA motifs or *cis* elements are shown in Figure 1. These motifs include, without limitation, glucocorticoid response motif(s), shear stress response motif(s), NFκB recognition motif(s), and AP1 motif(s). The locations of these and other motifs, discussed below, are diagramatically set forth in Figure 1.

As used herein, the term "*cis* elements capable of binding" refers to the ability of one or more of the described *cis* elements to specifically bind an agent. Such binding may be by any chemical, physical or biological interaction between the *cis* element and the agent, including, but not limited, to any covalent, steric, agostic, electronic and ionic interaction between the *cis* element and the agent. As used herein, the term "specifically binds" refers to the ability of the agent to bind to a specified *cis* element but not to wholly unrelated nucleic acid sequences. Regulatory region refers to the ability of a nucleic acid fragment, region or length to functionally perform a biological activity. The biological activity may be binding to the nucleic or specific DNA sequence. The biological activity may also modulate, enhance, inhibit or alter the transcription of a nearby coding region. The biological activity may be identified by a gel shift assay, in which binding to a nucleic acid fragment can be detected. Other methods of detecting the biological activity in a nucleic acid regulatory region are known in the art (*see Current Protocols in Molecular Biology*, for example).

Expression of the rat PRL gene is highly restricted to pituitary lactotroph cells and is induced by the cAMP-dependent protein kinase A pathway. At least one of the redundant pituitary specific elements (PRL-FP111) of the proximal rat PRL promotor is required for this protein kinase A effect (Rajnarayan *et al., Molecular Endocrinology 4*: 502-512 (1995), herein incorporated by reference). A sequence corresponding to an upstream motif or *cis* element characteristic of PRL-FP111 is set forth in Figure 1 at residues 370-388 and 4491-4502, respectively. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of molecules that bind the PRL-FP111 upstream motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

A consensus sequence (GR/PR), recognized by both the glucocorticoid receptor of rat liver and the progesterone receptor from rabbit uterus, has been reported to be involved in glucocorticoid and progesterone-dependent gene expression (Von der Ahe *et al.*, *Nature 313*:

706-709 (1985), herein incorporated by reference). A sequence corresponding to a GC/PR upstream motif or *cis* element is set forth in Figure 1 at residues 433-445. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of glucocorticoid or progesterone or their homologues, including, but not limited to, the concentration of glucocorticoid or progesterone or their homologues bound to an GC/PR upstream motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

Shear stress motif (SSRE) or *cis* element has been identified in a number of genes including platelet-derived growth factor B chain, tissue plasminogen activator (tPA), ICAM-1 and TGF-β1 (Resnick *et al., Proc. Natl. Acad. Sci. (USA) 80:* 4591-4595 (1993), herein incorporated by reference). Transcription of these genes has been associated with humoral stimuli such as cytokines and bacterial products as well as hemodynamic stress forces. Sequences corresponding to a upstream shear stress motif or *cis* element are set forth in Figure 1 at residues 446-451, 1288-1293, 3597-3602, 4771-4776, and 5240-5245, respectively. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of molecules capable of binding the shear stress motif. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

A consensus sequence for a glucocorticoid response upstream motif (GRE) or *cis* element has been characterized (Beato, *Cell 56:* 335-344 (1989); Becker *et al., Nature 324:* 686-688 (1986), herein incorporated by reference; Sakai *et al., Genes and Development 2:* 1144-1154 (1988), herein incorporated by reference). Genes containing this upstream motif or *cis* element are regulated by glucocorticoids, progesterone, androgens and mineral corticoids (Beato, *Cell 56:* 335-344 (1989)). Sequences corresponding to glucocorticoid response upstream motif or *cis* element are set forth in Figure 1 at residues 574-600, 1042-1056, 2444-2468, 2442-2269, 3536-3563, 4574-4593, 4595-4614, 4851-4865, 4844-4864, 5079-5084, and 5083-5111, respectively. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of molecules capable of binding a glucocorticoid response upstream motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

A sequence specific binding site (CBE) for the wild type nuclear phosphoprotein, p53, has been identified and appears to be associated with replication origins (Kern *et al. Science 252*: 1708-1711 (1991), herein incorporated by reference). A sequence corresponding to an CBE upstream motif or *cis* element is set forth in Figure 1 at residues 735-746. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of p53 or its homologues, including, but not limited to, the concentration of p53 or its homologues bound to an CBE upstream motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

Nuclear factor ets-like (NFE), a transcriptional activator that facilitates p50 and c-Rel-dependent IgH 3' enhancer activity has been shown to bind to an NFE site in the Rel-dependent IgH 3' enhancer (Linderson *et al.*, *European J. Immunology 27*: 468-475 (1997), herein incorporated by reference). A sequence corresponding to an NFE upstream motif or *cis* element is set forth in Figure 1 at residues 774-795. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of nuclear factors or their homologues, including, but not limited to, the concentration of nuclear factors or their homologues bound to an NFE upstream motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

An upstream motif or *cis* element (KTF.1-CS) for a control element 3' to the human keratin 1 gene that regulates cell type and differentiation-specific expression has been identified (Huff *et al.*, *J. Biological Chemistry 268:* 377-384 (1993), herein incorporated by reference). A sequence corresponding to an upstream motif or *cis* element characteristic of KTF.1-CS is set forth in Figure 1 at residues 843-854. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of KTF.1-CS or its homologues, including, but not limited to, the concentration of KTF.1-CS or its homologues bound to a KTF.1-CS upstream motif or cis element Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

A progesterone responsive element (PRE) that maps to the far upstream steroid dependent DNase hypersensitive site of chicken lysozyme chromatin has been characterized (Hecht *et al.*, *EMBO J. 7*: 2063-2073 (1988), herein incorporated by reference). The element confers hormonal regulation to a heterologous promoter and is composed of a cluster of progesterone

receptor binding sites. A sequence corresponding to an upstream motif or *cis* element characteristic of PRE is set forth in Figure 1 at residues 987-1026. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of molecules capable of binding a progesterone responsive PRE upstream motif or *cis* element. Such agents may be useful in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

A sequence (ETF-EGFR) has been characterized which serves as a motif for a *trans*-active transcription factor that regulates expression of the epidermal growth factor receptor (Regec *et al.*, *Blood 85*:2711-2719 (1995), herein incorporated by reference). A sequence corresponding to an ETF-EGFR upstream motif or *cis* element is set forth in Figure 1 at residues 1373-1388. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of nuclear factors or their homologues, including, but not limited to, the concentration of nuclear factors or their homologues bound to an ETF-EGFR upstream motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

A common trans-acting factor (SRE-cFos) has been shown to regulate skeletal and cardiac alpha-Actin gene transcription in muscle (Muscat *et al.*, *Molecular and Cellular Biology 10*: 4120-4133 (1988), herein incorporated by reference). A sequence corresponding to an SRE-cFos upstream motif or *cis* element is set forth in Figure 1 at residues 1447-1456. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of nuclear factors or their homologues, including, but not limited to, the concentration of nuclear factors or their homologues bound to an SRE-cFos upstream motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

Alu repetitive elements are unique to primates and are interspersed within the human genome with an average spacing of 4Kb. While some Alu sequences are actively transcribed by polymerase III, normal transcripts may also contain Alu-derived sequences in 5' or 3' untranslated regions (Jurka and Mikahanljaia, *J. Mol. Evolution 32*: 105-121 (1991), herein incorporated by reference, Claveria and Makalowski, *Nature 371*: 751-752 (1994), herein incorporated by reference). A sequence corresponding to an Alu upstream motif or *cis* element is

set forth in Figure 1 at residues 1331-1550. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of nuclear factors or their homologues, including, but not limited to, the concentration of nuclear factors or their homologues bound to an Alu upstream motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

A consensus sequence for a vitellogenin gene-binding protein (VBP) upstream motif or *cis* element has been characterized (Iyer *et al., Molecular and Cellular Biology 11:* 4863-4875 (1991), herein incorporated by reference). Expression of the VBP gene commences early in liver ontogeny and is not subject to circadian control. A sequence corresponding to an upstream motif or *cis* element capable of binding VBP is set forth in Figure 1 at residues 1786-1797. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of VBP or its homologues, including, but not limited to, the concentration of VBP or its homologues bound to an VBP upstream motif or cis element Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

A structural motif (Malt-CS) or *cis* element involved in the activation of all promoters of the maltose operons in *Escherichia coli* and *Klebsiella pneumoniae* has been characterized (Vidal-Ingigliardi *et al., J. Mol. Biol. 218:* 323-334 (1991), herein incorporated by reference). A sequence corresponding to a upstream Malt-CS motif or *cis* element is set forth in Figure 1 at residues 1832-1841. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of molecules capable of binding the upstream Malt-CS motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

A consensus sequence for an estrogen receptor upstream motif or *cis* element has been characterized (ERE) (Forman *et al., Mol. Endocrinology 4:* 1293-1301 (1990), herein incorporated by reference; de Verneuil *et al., Nucleic Acid Res. 18:* 4489-4497 (1990), herein incorporated by reference; Gaub *et al., Cell 63:* 1267-1276 (1990), herein incorporated by reference. A sequence corresponding to half an upstream motif or *cis* element capable of binding estrogen receptor is set forth in Figure 1 at residues 2166-2195, 3413-3429, and 3892-3896, respectively. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or

concentration, of the estrogen receptor or its homologues bound to an upstream motif or cis element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

Certain protein-binding sites (NF-mutagen) in Ig gene enhancers which determine transcriptional activity and inducibility have been shown to interact with nuclear factors (Lenardo *et al.*, *Science 236*: 1573-1577 (1987), herein incorporated by reference). A sequence corresponding to an NF-mutagen upstream motif or *cis* element is set forth in Figure 1 at residues 2329-2338. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of nuclear factors or their homologues, including, but not limited to, the concentration of nuclear factors or their homologues bound to an NF-mutagen upstream motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

A consensus sequence for a transcriptional repressor of c-myc (myc-PRF) upstream motif or *cis* element has been identified (Kakkis *et al., Nature 339:* 718-719 (1989), herein incorporated by reference). Myc-PRF interacts with another widely distributed protein, myc-CF1 (common factor 1), which binds nearby and this association may be important in myc-PRF repression. A sequence corresponding to an upstream motif or *cis* element capable of binding myc-PRF is set forth in Figure 1 at residues 2403-2416. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of myc-PRF or its homologues, including, but not limited to, the concentration of myc-PRF or its homologues bound to an myc-PRF upstream motif or *cis* element Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

Human transcription factor activator protein 2 (AP2) is a transcription factor that has been shown to bind to Sp1, nuclear factor 1 (NF1) and simian virus 40 transplantation (SV40 T) antigen binding sites. It is developmentally regulated (Williams and Tijan, *Gene Dev. 5:* 670-682 (1991), herein incorporated by reference; Mitchell *et al., Genes Dev. 5:* 105-119 (1991), herein incorporated by reference; Coutois *et al., Nucleic Acid Research 18:* 57-64 (1990), herein incorporated by reference; Comb *et al., Nucleic Acid Research 18:* 3975-3982 (1990), herein incorporated by reference; Winings *et al., Nucleic Acid Research 19:* 3709-3714 (1991), herein incorporated by reference). Sequences corresponding to an upstream motif or *cis* element capable of binding AP2 are set forth in Figure 1 at residues 2520-2535, and 5170-5187,

respectively. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of AP2 or its homologues, including, but not limited to, the concentration of AP2 or its homologues bound to an upstream motif or cis element. Such agents may be useful in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

Drosophila RNA polymerase II heat shock transcription factor (HSTF) is a transcription factor that has been shown to be required for active transcription of an hsp 70 gene (Parker and Topol, *Cell 37:* 273-283 (1984), herein incorporated by reference). Sequences corresponding to an upstream motif or *cis* element capable of binding HSTF are set forth in Figure 1 at residues 2622-2635, and 5105-5132. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of HSTF or its homologues, including, but not limited to, the concentration of HSTF or its homologues bound to an HSTF upstream motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

A sequence corresponding to an upstream motif or *cis* element characteristic of SBF is set forth in Figure 1 at residues 2733-2743 (Shore *et al., EMBO J. 6:* 461-467 (1987), herein incorporated by reference). In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of molecules that bind the SBF upstream motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

An NF1 motif or *cis* element has been identified which recognizes a family of at least six proteins (Courtois, *et al., Nucleic Acid Res. 18:* 57-64 (1990), herein incorporated by reference; Mul *et al., J. Virol. 64:* 5510-5518 (1990), herein incorporated by reference; Rossi *et al., Cell 52:* 405-414 (1988), herein incorporated by reference; Gounari *et al., EMBO J. 10:* 559-566 (1990), herein incorporated by reference; Goyal *et al., Mol. Cell Biol. 10:* 1041-1048 (1990); herein incorporated by reference; Mermond *et al., Nature 332:* 557-561 (1988), herein incorporated by reference; Gronostajski *et al., Molecular and Cellular Biology 5:* 964-971 (1985), herein incorporated by reference; Hennighausen *et al., EMBO J. 5:* 1367-1371 (1986), herein incorporated by reference; Chodosh *et al., Cell 53:* 11-24 (1988), herein incorporated by reference). The NF1 protein will bind to an NF1 motif or *cis* element either as a dimer (if the

motif is palindromic) or as an single molecule (if the motif is not palindromic). The NF1 protein is induced by TGFβ (Faisst and Meyer, *Nucleic Acid Research 20:* 3-26 (1992), herein incorporated by reference). Sequences corresponding to an upstream motif or *cis* element capable of binding NF1 are set forth in Figure 1 at residues 2923-2938, 4143-4167, and 4886-4900, respectively. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of NF1 or its homologues, including, but not limited to, the concentration of NF1 or its homologues bound to an upstream motif or cis element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

Conserved regulatory sequences (NF-MHCIIA/B) of a rabbit major histocompatability complex (MHC) class II gene are responsible for binding two distinct nuclear factors NF-MHCIIA and NF-MHCIIB and are believed to be involved in the regulation of coordinate expression of the class II genes -- eg. MHC class II gene in B lymphocytes (Sittisombut *Molecular and Cellular Biology 5*: 2034-2041 (1988), herein incorporated by reference). A sequence corresponding to an NF-MHCIIA/B upstream motif or *cis* element is set forth in Figure 1 at residues 2936-2944. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of NF-MHCIIA or NF-MHCIIB or their homologues, including, but not limited to, the concentration of NF-MHCIIA or NF-MHCIIB or their homologues bound to an NF-MHCIIA/B upstream motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

PEA 1 binding motifs or cis elements have been identified (Piette and Yaniv, *EMBO J. 5*: 1331-1337 (1987), herein incorporated by reference). The PEA1 protein is a transcription factor that is reported to bind to both the polyoma virus and c-*fos* enhancers. A sequence corresponding to an upstream motif or *cis* element capable of binding PEA1 is set forth in Figure 1 at residues 3285-3298. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of PEA1 or its homologues, including, but not limited to, the concentration of PEA1 or its homologues bound to an upstream motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

A conserved cis-acting regulatory element (ICS) has been shown to bind trans-acting constituitive nuclear factors present in lymphocytes and fibroblasts which are involved in the interferon (IFN)-mediated transcriptional enhancement of MHC class I and other genes (Shirayoshi *et al., Proc. Natl. Acad. Sci. (USA) 85*: 5884-5888 (1988), herein incorporated by reference). A sequence corresponding to an ICS upstream motif or *cis* element is set forth in Figure 1 at residues 3688-3699. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of nuclear factors or their homologues, including, but not limited to, the concentration of nuclear factors or their homologues bound to an ICS upstream motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

A consensus sequence for an ISGF2 upstream motif or *cis* element has been characterized (Iman *et al., Nucleic Acids Res. 18:* 6573-6580 (1990), herein incorporated by reference; Harada *et al., Cell 63:* 303-312 (1990), herein incorporated by reference; Yu-Lee *et al., Mol. Cell Biol. 10:* 3087-3094 (1990), herein incorporated by reference; Pine *et al., Mol. Cell Biol. 10:* 32448-2457 (1990), herein incorporated by reference). ISGF2 is induced by interferon α and γ, prolactin and virus infections. A sequence corresponding to an upstream motif or *cis* element capable of binding ISGF2 is set forth in Figure 1 at residues 4170-4179. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of ISGF2 or its homologues, including, but not limited to, the concentration of ISGF2 or its homologues bound to an upstream motif or cis element Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

A sequence corresponding to an upstream motif or *cis* element capable of binding zinc is set forth in Figure 1 at residues 4285-4292. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of zinc. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

A sequence corresponding to an upstream motif or *cis* element characteristic of CAP/CRP-galO is set forth in Figure 1 at residues 4379-4404 (Taniguchi *et al., Proc. Natl. Acad. Sci (USA) 76:* 5090-5094 (1979), herein incorporated by reference). In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents

capable of altering the biochemical properties or concentration of molecules that bind the CAP/CRP-galO upstream motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

Human transcription factor activator protein 1 (AP1) is a transcription factor that has been shown to regulate genes which are highly expressed in transformed cells such as stromelysin, c-*fos*, α_1-anti-trypsin and collagenase (Gutman and Wasylyk, *EMBO J. 9.7:* 2241-2246 (1990), herein incorporated by reference; Martin *et al., Proc. Natl. Acad. Sci. USA 85:* 5839-5843 (1988), herein incorporated by reference; Jones *et al., Genes and Development 2:* 267-281 (1988), herein incorporated by reference; Faisst and Meyer, *Nucleic Acid Research 20:* 3-26 (1992), herein incorporated by reference; Kim *et al., Molecular and Cellular Biology 10:* 1492-1497 (1990), herein incorporated by reference: Baumhueter *et al., EMBO J. 7:* 2485-2493 (1988), herein incorporated by reference). The AP1 transcription factor has been associated with genes that are activated by 12-O-tetradecanolyphorbol-13-acetate (TPA) (Gutman and Wasylyk, *EMBO J.7:* 2241-2246 (1990)). Sequences corresponding to an upstream motif or *cis* element capable of binding AP1 are set forth in Figure 1 at residues 4428-4434 and 4627-4639, respectively. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of AP1 or its homologues, including, but not limited to, the concentration of AP1 or its homologues bound to an upstream motif or cis element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

The sex-determining region of the Y chromosome gene, sry, is expressed in the fetal mouse for a brief period, just prior to testis differentiation. SRY is a DNA binding protein known to bind to a CACA-rich region in the sry gene (Vriz *et al., Biochemistry and Molecular Biology International 37*: 1137-1146 (1995), herein incorporated by reference). A sequence corresponding to an upstream motif or *cis* element capable of binding SRY is set forth in Figure 1 at residues 4625-4634. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of SRY or its homologues, including, but not limited to, the concentration of SRY or its homologues bound to an upstream motif or *cis* element. Such agents may be useful in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

A sequence corresponding to an upstream motif or *cis* element characteristic of GC2-GH is set forth in Figure 1 at residues 4689-4711 (West *et al., Molecular and Cellular Biology 7:* 1193-1197 (1987), herein incorporated by reference). In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of GC2-GH or its homologues, including, but not limited to, the concentration of GC2-GH or its homologues bound to an upstream motif or cis element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

PEA 3 binding motifs or cis elements have been identified (Martin *et al., Proc. Natl. Acad. Sci. (USA) 85:* 5839-5843 (1988), herein incorporated by reference; Gutman and Wasylyk, *EMBO J. 7:* 2241-2246 (1990), herein incorporated by reference). The PEA3 protein is a transcription factor that is reported to interact with AP1 like proteins (Martin *et al., Proc. Natl. Acad. Sci. (USA) 85:* 5839-5843 (1988), herein incorporated by reference). Sequences corresponding to an upstream motif or *cis* element capable of binding PEA3 is set forth in Figure 1 at residues 4765-4769. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of PEA3 or its homologues, including, but not limited to, the concentration of PEA3 or its homologues bound to an upstream motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

Mammalian interspersed repetitive (MIR) is an element involved in the coding and processing sequences of mammalian genes. The MIR element is at least 260 bp in length and numbers about 10^5 copies within the mammalian genome (Murnane *et al.*, *Nucleic Acids Research 15*: 2837-2839 (1995), herein incorporated by reference). A sequence corresponding to an MIR upstream motif or *cis* element is set forth in Figure 1 at residues 4759-4954. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of nuclear factors or their homologues, including, but not limited to, the concentration of nuclear factors or their homologues bound to an MIR upstream motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

Normal liver and differentiated hepatoma cell lines contain a hepatocyte-specific nuclear factor (HNF-1) which binds cis-acting element sequences within the promoters of the alpha and

beta chains of fibrinogen and alpha 1-antitrypsin (Baumhueter *et al., EMBO J. 8*: 2485-2493, herein incorporated by reference). A sequence corresponding to an HNF-1 upstream motif or *cis* element is set forth in Figure 1 at residues 4923-4941. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of HNF-1 or its homologues, including, but not limited to, the concentration of HNF-1 or its homologues bound to an HNF-1 upstream motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

A number of *cis* elements or upstream motifs have been associated with gene regulation by steroid and thyroid hormones (e.g. glucocorticoid and estrogen)(Beato, *Cell 56:* 335-344 (1989), herein incorporated by reference; Brent *et al., Molecular Endocrinology 89:*1996-2000 (1989), herein incorporated by reference; Glass *et al., Cell 54:* 313-323 (1988), herein incorporated by reference; Evans, *Science 240:* 889-895 (1988), herein incorporated by reference).

A consensus sequence for a thyroid receptor upstream motif or *cis* element (TRE) has been characterized (Beato, *Cell 56:* 335-344 (1989), herein incorporated by reference). A sequence corresponding to a thyroid receptor upstream motif or *cis* element is set forth in Figure 1 at residues 5151-5156. Thyroid hormones are capable of regulating genes containing a thyroid receptor upstream motif or *cis* element (Glass *et al., Cell 54:* 313-323 (1988), herein incorporated by reference). Thyroid hormones can negatively regulate TIGR. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of molecules capable of binding a thyroid receptor upstream motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

NFκB is a transcription factor that is reportedly associated with a number of biological processes including T-cell activation and cytokine regulation (Lenardo *et al., Cell 58:* 227-229 (1989), herein incorporated by reference). A consensus upstream motif or *cis* element capable of binding NFκB has been reported (Lenardo *et al., Cell 58:* 227-229 (1989)). Sequences corresponding to an upstream motif or *cis* element capable of binding NFκB are set forth in Figure 1 at residues 5166-5175. In accordance with the embodiments of the present invention, transcription of TIGR molecules can be effected by agents capable of altering the biochemical properties or concentration of NFκB or its homologues, including, but not limited to, the concentration of NFκB or its homologues bound to an upstream motif or *cis* element. Such agents can be used in the study of glaucoma pathogenesis. In another embodiment, such agents

can also be used in the study of glaucoma prognosis. In another embodiment such agents can be used in the treatment of glaucoma.

Illustrative Uses of the Nucleic Acids of the Invention

Where one or more of the agents is a nucleic acid molecule, such nucleic acid molecule may be sense, antisense or triplex oligonucleotides corresponding to any part of the TIGR promoter, TIGR cDNA, TIGR intron, TIGR exon or TIGR gene. In some embodiments these nucleic acids may be about 20 bases in length, as for example, SEQ. ID NO: 6-25 or 33. In some circmstances, the nucleic acids may be only about 8 bases in length. Short nucleic acids may be particularly useful in hybridization to immobilized nucleic acids in order to determine the presence of specific sequences, such as by the known methods of sequencing by hybridization.

The TIGR promoter, or fragment thereof, of the present invention may be cloned into a suitable vector and utilized to promote the expression of a marker gene (e.g. firefly luciferase (de Wet, *Mol. Cell Biol. 7:* 725-737 (1987), herein incorporated by reference) or GUS (Jefferson *et al., EMBO J. 6:* 3901-3907 (1987), herein incorporated by reference)). In another embodiment of the present invention, a TIGR promoter may be cloned into a suitable vector and utilized to promote the expression of a TIGR gene in a suitable eukaryotic or prokaryotic host cell (e.g. human trabecular cell, chinese hamster cell, *E. coli*). In another embodiment of the present invention, a TIGR promoter may be cloned into a suitable vector and utilized to promote the expression of a homologous or heterologous gene in a suitable eukaryotic or prokaryotic host cells (e.g. human trabecular cell lines, chinese hamster cells, *E. coli*).

Practitioners are familiar with the standard resource materials which describe specific conditions and procedures for the construction, manipulation and isolation of macromolecules (e.g., DNA molecules, plasmids, etc.), generation of recombinant organisms and the screening and isolating of clones, (see for example, Sambrook *et al.,* In *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press (1989), herein incorporated by reference in its entirety; Old and Primrose, In Principles of Gene Manipulation: An Introduction To Genetic Engineering, Blackwell (1994), herein incorporated by reference).

The TIGR promoter, or any portion thereof, or an about 10 to about 500 bases fragment thereof, of the present invention may be used in a gel-retardation or band shift assay (Old and Primrose, In Principles of Gene Manipulation: An Introduction To Genetic Engineering, Blackwell (1994)). Nucleic acids or fragments comprising any of the *cis* elements identified in the present invention may be used in a gel-retardation or band shift assay to isolate proteins capable of binding the *cis* element. Suitable DNA fragments or molecules comprise or consist of one or more of the following: sequences corresponding to an upstream motif or *cis* element

characteristic of PRL-FP111 as set forth in Figure 1 at residues 370-388, and 4491-4502, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding GR/PR as set forth in Figure 1 at residues 433-445, sequences corresponding to an upstream shear stress motif or *cis* element as set forth in Figure 1 at residues 446-451, 1288-1293, 3597-3602, 4771-4776, and 5240-5245, respectively, sequences corresponding to glucocorticoid response upstream motif or *cis* element as set forth in Figure 1 at residues 574-600, 1042-1056, 2444-2468, 2442-2269, 3536-3563, 4574-4593, 4595-4614, 4851-4865, 4844-4864, 5079-5084, 5083-5111, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding CBE as set forth in Figure 1 at residues 735-746, a sequence corresponding to an upstream motif or *cis* element capable of binding NFE as set forth in Figure 1 at residues 774-795, a sequence corresponding to an upstream motif or *cis* element capable of binding KTF.1-CS as set forth in Figure 1 at residues 843-854, a sequence corresponding to an upstream motif or *cis* element capable of binding PRE is set forth in Figure 1 at residues 987-1026, a sequence corresponding to an upstream motif or *cis* element capable of binding ETF-EGFR as set forth in Figure 1 at residues 1373-1388, a sequence corresponding to an upstream motif or *cis* element capable of binding SRE-cFos as set forth in Figure 1 at residues 1447-1456, a sequence corresponding to an upstream motif or *cis* element capable of binding Alu as set forth in Figure 1 at residues 1331-1550, a sequence corresponding to an upstream motif or *cis* element capable of binding VBP as set forth in Figure 1 at residues 1786-1797, a sequence corresponding to an upstream motif or *cis* element capable of binding Malt-CS as set forth in Figure 1 at residues 1832-1841, sequences corresponding to an upstream motif or *cis* element capable of binding ERE as set forth in Figure 1 at residues 2167-2195, 3413-3429, and 3892-3896, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding NF-mutagen as set forth in Figure 1 at residues 2329-2338, a sequence corresponding to an upstream motif or *cis* element capable of binding myc-PRF as set forth in Figure 1 at residues 2403-2416, sequences corresponding to an upstream motif or *cis* element capable of binding AP2 as set forth in Figure 1 at residues 2520-2535 and 5170-5187, respectively, sequences corresponding to an upstream motif or *cis* element capable of binding HSTF as set forth in Figure 1 at residues 2622-2635, and 5105-5132, respectively, a sequence corresponding to an upstream motif or *cis* element characteristic of SBF as set forth in Figure 1 at residues 2733-2743, sequences corresponding to an upstream motif or *cis* element capable of binding NF-1 as set forth in Figure 1 at residues 2923-2938, 4144-4157, and 4887-4900, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding NF-MHCIIA/B as set forth in Figure 1 at residues 2936-2944, a sequence corresponding to an upstream motif or *cis* element capable of binding PEA1 as set forth in Figure 1 at residues 3285-3298, a sequence corresponding to an upstream motif or *cis* element capable of binding ICS as set forth in Figure 1 at residues 3688-3699, a sequence

corresponding to an upstream motif or *cis* element capable of binding ISGF2 as set forth in Figure 1 at residues 4170-4179, a sequence corresponding to an upstream motif or *cis* element capable of binding zinc as set forth in Figure 1 at residues 4285-4293, a sequence corresponding to an upstream motif or *cis* element characteristic of CAP/CRP-galO as set forth in Figure 1 at residues 4379-4404, sequences corresponding to an upstream motif or *cis* element capable of binding AP1 as set forth in Figure 1 at residues 4428-4434, and 4627-4639, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding SRY as set forth in Figure 1 at residues 4625-4634, a sequence corresponding to an upstream motif or *cis* element characteristic of GC2 as set forth in Figure 1 at residues 4678-4711, a sequence corresponding to an upstream motif or *cis* element capable of binding PEA3 as set forth in Figure 1 at residues 4765-4769, a sequence corresponding to an upstream motif or *cis* element capable of MIR as set forth in Figure 1 at residues 4759-4954, a sequence corresponding to an upstream motif or *cis* element capable of binding NF-HNF-1 as set forth in Figure 1 at residues 4923-4941, a sequence corresponding to a thyroid receptor upstream motif or *cis* element as set forth in Figure 1 at residues 5151-5156, and a sequence corresponding to an upstream motif or *cis* element capable of binding NFκB as set forth in Figure 1 at residues 5166-5175.

A preferred class of agents of the present invention comprises nucleic acid molecules encompassing all or a fragment of the "TIGR promoter" or 5' flanking gene sequences. As used herein, the terms "TIGR promoter" or "promoter" is used in an expansive sense to refer to the regulatory sequence(s) that control mRNA production. Thus, TIGR promoter sequences can be identified by those sequences that functionally effect the intiation, rate, or amount of transcription of the TIGR gene product mRNA. Such sequences include RNA polymerase binding sites, glucocorticoid response elements, enhancers, etc. These sequences may preferably be found within the specifically disclosed 5' upstream region sequences disclosed here, and most preferably within an about 500 base region 5' to the start of transcription or within an about 300 base region 5'of the transcription start site. However, other genomic sequences may be a TIGR promoter. Methods known in the art to identify distant promoter elements can be used with the disclosed sequences and nucleic acids to identify and define these distant TIGR promoter sequences. Such TIGR molecules may be used to diagnose the presence of glaucoma and the severity of or susceptibility to glaucoma. Such molecules may be either DNA or RNA.

A functional regulatory region of the TIGR gene may be a TIGR promoter ~~sequence.~~ It may also include transcription enhancer sites and transcription inhibitor sites or binding sites for a number of known proteins or molecules demonstrated as effecting transcription. A number of regulatory elements are discussed below, and the equivalent of those activities can represent the functional regulatory region of the TIGR gene. The methods for identifying and detecting the activity and function of these regulatory regions are known in the art.

Fragment nucleic acid molecules may encode significant portion(s) of, or indeed most of, SEQ ID NO: 1 or SEQ ID NO: 3 or SEQ ID NO: 4 or SEQ ID NO: 5. Alternatively, the fragments may comprise smaller oligonucleotides (having from about 15 to about 250 nucleotide residues, and more preferably, about 15 to about 30 nucleotide residues.). Such oligonucleotides include SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25.

Alternatively such oligonucleotides may derive from either the TIGR promoter, TIGR introns, TIGR exons, TIGR cDNA and TIGR downstream sequences comprise or consist of one or more of the following: sequences corresponding to an upstream motif or *cis* element characteristic of PRL-FP111 as set forth in Figure 1 at residues 370-388, and 4491-4502, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding GR/PR as set forth in Figure 1 at residues 433-445, sequences corresponding to an upstream shear stress motif or *cis* element as set forth in Figure 1 at residues 446-451, 1288-1293, 3597-3602, 4771-4776, and 5240-5245, respectively, sequences corresponding to glucocorticoid response upstream motif or *cis* element as set forth in Figure 1 at residues 574-600, 1042-1056, 2444-2468, 2442-2269, 3536-3563, 4574-4593, 4595-4614, 4851-4865, 4844-4864, 5079-5084, 5083-5111, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding CBE as set forth in Figure 1 at residues 735-746, a sequence corresponding to an upstream motif or *cis* element capable of binding NFE as set forth in Figure 1 at residues 774-795, a sequence corresponding to an upstream motif or *cis* element capable of binding KTF.1-CS as set forth in Figure 1 at residues 843-854, a sequence corresponding to an upstream motif or *cis* element capable of binding PRE is set forth in Figure 1 at residues 987-1026, a sequence corresponding to an upstream motif or *cis* element capable of binding ETF-EGFR as set forth in Figure 1 at residues 1373-1388, a sequence corresponding to an upstream motif or *cis* element capable of binding SRE-cFos as set forth in Figure 1 at residues 1447-1456, a sequence corresponding to an upstream motif or *cis* element capable of binding Alu as set forth in Figure 1 at residues 1331-1550, a sequence corresponding to an upstream motif or *cis* element capable of binding VBP as set forth in Figure 1 at residues 1786-1797, a sequence corresponding to an upstream motif or *cis* element capable of binding Malt-CS as set forth in Figure 1 at residues 1832-1841, sequences corresponding to an upstream motif or *cis* element capable of binding ERE as set forth in Figure 1 at residues 2167-2195, 3413-3429, and 3892-3896, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding NF-mutagen as set forth in Figure 1 at residues 2329-2338, a sequence corresponding to an upstream motif or *cis* element capable of binding myc-PRF as set forth in Figure 1 at residues 2403-2416, sequences

corresponding to an upstream motif or *cis* element capable of binding AP2 as set forth in Figure 1 at residues 2520-2535 and 5170-5187, respectively, sequences corresponding to an upstream motif or *cis* element capable of binding HSTF as set forth in Figure 1 at residues 2622-2635, and 5105-5132, respectively, a sequence corresponding to an upstream motif or *cis* element characteristic of SBF as set forth in Figure 1 at residues 2733-2743, sequences corresponding to an upstream motif or *cis* element capable of binding NF-1 as set forth in Figure 1 at residues 2923-2938, 4144-4157, and 4887-4900, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding NF-MHCIIA/B as set forth in Figure 1 at residues 2936-2944, a sequence corresponding to an upstream motif or *cis* element capable of binding PEA1 as set forth in Figure 1 at residues 3285-3298, a sequence corresponding to an upstream motif or *cis* element capable of binding ICS as set forth in Figure 1 at residues 3688-3699, a sequence corresponding to an upstream motif or *cis* element capable of binding ISGF2 as set forth in Figure 1 at residues 4170-4179, a sequence corresponding to an upstream motif or *cis* element capable of binding zinc as set forth in Figure 1 at residues 4285-4293, a sequence corresponding to an upstream motif or *cis* element characteristic of CAP/CRP-galO as set forth in Figure 1 at residues 4379-4404, sequences corresponding to an upstream motif or *cis* element capable of binding AP1 as set forth in Figure 1 at residues 4428-4434, and 4627-4639, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding SRY as set forth in Figure 1 at residues 4625-4634, a sequence corresponding to an upstream motif or *cis* element characteristic of GC2 as set forth in Figure 1 at residues 4678-4711, a sequence corresponding to an upstream motif or *cis* element capable of binding PEA3 as set forth in Figure 1 at residues 4765-4769, a sequence corresponding to an upstream motif or *cis* element capable of MIR as set forth in Figure 1 at residues 4759-4954, a sequence corresponding to an upstream motif or *cis* element capable of binding NF-HNF-1 as set forth in Figure 1 at residues 4923-4941, a sequence corresponding to a thyroid receptor upstream motif or *cis* element as set forth in Figure 1 at residues 5151-5156, and a sequence corresponding to an upstream motif or *cis* element capable of binding NFκB as set forth in Figure 1 at residues 5166-5175. For such purpose, the oligonucleotides must be capable of specifically hybridizing to a nucleic acid molecule genetically or physically linked to the TIGR gene. As used herein, the term "linked" refers to genetically, physically or operably linked.

As used herein, two nucleic acid molecules are said to be capable of specifically hybridizing to one another if the two molecules are capable of forming an anti-parallel, double-stranded nucleic acid structure, whereas they are unable to form a double-stranded structure when incubated with a non-TIGR nucleic acid molecule. A nucleic acid molecule is said to be the "complement" of another nucleic acid molecule if they exhibit complete complementarity. As used herein, molecules are said to exhibit "complete complementarity" when every nucleotide of

one of the molecules is complementary to a nucleotide of the other. Two molecules are said to be "minimally complementary" if they can hybridize to one another with sufficient stability to permit them to remain annealed to one another under at least conventional "low-stringency" conditions. Similarly, the molecules are said to be "complementary" if they can hybridize to one another with sufficient stability to permit them to remain annealed to one another under conventional "high-stringency" conditions. Conventional stringency conditions are described by Sambrook, J., *et al.*, (In: *Molecular Cloning, a Laboratory Manual, 2nd Edition, Cold Spring Harbor Press*, Cold Spring Harbor, New York (1989)), and by Haymes, B.D., *et al.* (In: *Nucleic Acid Hybridization, A Practical Approach*, IRL Press, Washington, DC (1985)), both herein incorporated by reference). Departures from complete complementarity are therefore permissible, as long as such departures do not completely preclude the capacity of the molecules to form a double-stranded structure. Thus, in order for an oligonucleotide to serve as a primer it need only be sufficiently complementary in sequence to be able to form a stable double-stranded structure under the particular solvent and salt concentrations employed.

Apart from their diagnostic or prognostic uses, such oligonucleotides may be employed to obtain other TIGR nucleic acid molecules. Such molecules include the TIGR-encoding nucleic acid molecule of non-human animals (particularly, cats, monkeys, rodents and dogs), fragments thereof, as well as their promoters and flanking sequences. Such molecules can be readily obtained by using the above-described primers to screen cDNA or genomic libraries obtained from non-human species. Methods for forming such libraries are well known in the art. Such analogs may differ in their nucleotide sequences from that of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, or from molecules consisting of sequences corresponding to an upstream motif or *cis* element characteristic of PRL-FP111 as set forth in Figure 1 at residues 370-388, and 4491-4502, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding GR/PR as set forth in Figure 1 at residues 433-445, sequences corresponding to an upstream shear stress motif or *cis* element as set forth in Figure 1 at residues 446-451, 1288-1293, 3597-3602, 4771-4776, and 5240-5245, respectively, sequences corresponding to glucocorticoid response upstream motif or *cis* element as set forth in Figure 1 at residues 574-600, 1042-1056, 2444-2468, 2442-2269, 3536-3563, 4574-4593, 4595-4614, 4851-4865, 4844-4864, 5079-5084, 5083-5111, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding CBE as set forth in Figure 1 at residues 735-746, a sequence corresponding to an upstream motif or *cis* element capable of binding NFE as set forth in Figure 1 at residues 774-

795, a sequence corresponding to an upstream motif or *cis* element capable of binding KTF.1-CS as set forth in Figure 1 at residues 843-854, a sequence corresponding to an upstream motif or *cis* element capable of binding PRE is set forth in Figure 1 at residues 987-1026, a sequence corresponding to an upstream motif or *cis* element capable of binding ETF-EGFR as set forth in Figure 1 at residues 1373-1388, a sequence corresponding to an upstream motif or *cis* element capable of binding SRE-cFos as set forth in Figure 1 at residues 1447-1456, a sequence corresponding to an upstream motif or *cis* element capable of binding Alu as set forth in Figure 1 at residues 1331-1550, a sequence corresponding to an upstream motif or *cis* element capable of binding VBP as set forth in Figure 1 at residues 1786-1797, a sequence corresponding to an upstream motif or *cis* element capable of binding Malt-CS as set forth in Figure 1 at residues 1832-1841, sequences corresponding to an upstream motif or *cis* element capable of binding ERE as set forth in Figure 1 at residues 2167-2195, 3413-3429, and 3892-3896, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding NF-mutagen as set forth in Figure 1 at residues 2329-2338, a sequence corresponding to an upstream motif or *cis* element capable of binding myc-PRF as set forth in Figure 1 at residues 2403-2416, sequences corresponding to an upstream motif or *cis* element capable of binding AP2 as set forth in Figure 1 at residues 2520-2535 and 5170-5187, respectively, sequences corresponding to an upstream motif or *cis* element capable of binding HSTF as set forth in Figure 1 at residues 2622-2635, and 5105-5132, respectively, a sequence corresponding to an upstream motif or *cis* element characteristic of SBF as set forth in Figure 1 at residues 2733-2743, sequences corresponding to an upstream motif or *cis* element capable of binding NF-1 as set forth in Figure 1 at residues 2923-2938, 4144-4157, and 4887-4900, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding NF-MHCIIA/B as set forth in Figure 1 at residues 2936-2944, a sequence corresponding to an upstream motif or *cis* element capable of binding PEA1 as set forth in Figure 1 at residues 3285-3298, a sequence corresponding to an upstream motif or *cis* element capable of binding ICS as set forth in Figure 1 at residues 3688-3699, a sequence corresponding to an upstream motif or *cis* element capable of binding ISGF2 as set forth in Figure 1 at residues 4170-4179, a sequence corresponding to an upstream motif or *cis* element capable of binding zinc as set forth in Figure 1 at residues 4285-4293, a sequence corresponding to an upstream motif or *cis* element characteristic of CAP/CRP-galO as set forth in Figure 1 at residues 4379-4404, sequences corresponding to an upstream motif or *cis* element capable of binding AP1 as set forth in Figure 1 at residues 4428-4434, and 4627-4639, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding SRY as set forth in Figure 1 at residues 4625-4634, a sequence corresponding to an upstream motif or *cis* element characteristic of GC2 as set forth in Figure 1 at residues 4678-4711, a sequence corresponding to an upstream motif or *cis* element capable of binding PEA3 as set forth in Figure 1 at residues

4765-4769, a sequence corresponding to an upstream motif or *cis* element capable of MIR as set forth in Figure 1 at residues 4759-4954, a sequence corresponding to an upstream motif or *cis* element capable of binding NF-HNF-1 as set forth in Figure 1 at residues 4923-4941, a sequence corresponding to a thyroid receptor upstream motif or *cis* element as set forth in Figure 1 at residues 5151-5156, and a sequence corresponding to an upstream motif or *cis* element capable of binding NFκB as set forth in Figure 1 at residues 5166-5175 because complete complementarity is not needed for stable hybridization. The TIGR nucleic acid molecules of the present invention therefore also include molecules that, although capable of specifically hybridizing with TIGR nucleic acid molecules may lack "complete complementarity."

Any of a variety of methods may be used to obtain the above-described nucleic acid molecules (Elles, Methods in Molecular Medicine: Molecular Diagnosis of Genetic Diseases, Humana Press (1996), herein incorporated by reference). SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 33, sequences corresponding to an upstream motif or *cis* element characteristic of PRL-FP111 as set forth in Figure 1 at residues 370-388, and 4491-4502, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding GR/PR as set forth in Figure 1 at residues 433-445, sequences corresponding to an upstream shear stress motif or *cis* element as set forth in Figure 1 at residues 446-451, 1288-1293, 3597-3602, 4771-4776, and 5240-5245, respectively, sequences corresponding to glucocorticoid response upstream motif or *cis* element as set forth in Figure 1 at residues 574-600, 1042-1056, 2444-2468, 2442-2269, 3536-3563, 4574-4593, 4595-4614, 4851-4865, 4844-4864, 5079-5084, 5083-5111, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding CBE as set forth in Figure 1 at residues 735-746, a sequence corresponding to an upstream motif or *cis* element capable of binding NFE as set forth in Figure 1 at residues 774-795, a sequence corresponding to an upstream motif or *cis* element capable of binding KTF.1-CS as set forth in Figure 1 at residues 843-854, a sequence corresponding to an upstream motif or *cis* element capable of binding PRE is set forth in Figure 1 at residues 987-1026, a sequence corresponding to an upstream motif or *cis* element capable of binding ETF-EGFR as set forth in Figure 1 at residues 1373-1388, a sequence corresponding to an upstream motif or *cis* element capable of binding SRE-cFos as set forth in Figure 1 at residues 1447-1456, a sequence corresponding to an upstream motif or *cis* element capable of binding Alu as set forth in Figure 1 at residues 1331-1550, a sequence corresponding to an upstream motif or *cis* element capable of binding VBP as set forth in Figure 1 at residues 1786-1797, a sequence corresponding to an upstream motif or *cis*

element capable of binding Malt-CS as set forth in Figure 1 at residues 1832-1841, sequences corresponding to an upstream motif or *cis* element capable of binding ERE as set forth in Figure 1 at residues 2167-2195, 3413-3429, and 3892-3896, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding NF-mutagen as set forth in Figure 1 at residues 2329-2338, a sequence corresponding to an upstream motif or *cis* element capable of binding myc-PRF as set forth in Figure 1 at residues 2403-2416, sequences corresponding to an upstream motif or *cis* element capable of binding AP2 as set forth in Figure 1 at residues 2520-2535 and 5170-5187, respectively, sequences corresponding to an upstream motif or *cis* element capable of binding HSTF as set forth in Figure 1 at residues 2622-2635, and 5105-5132, respectively, a sequence corresponding to an upstream motif or *cis* element characteristic of SBF as set forth in Figure 1 at residues 2733-2743, sequences corresponding to an upstream motif or *cis* element capable of binding NF-1 as set forth in Figure 1 at residues 2923-2938, 4144-4157, and 4887-4900, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding NF-MHCIIA/B as set forth in Figure 1 at residues 2936-2944, a sequence corresponding to an upstream motif or *cis* element capable of binding PEA1 as set forth in Figure 1 at residues 3285-3298, a sequence corresponding to an upstream motif or *cis* element capable of binding ICS as set forth in Figure 1 at residues 3688-3699, a sequence corresponding to an upstream motif or *cis* element capable of binding ISGF2 as set forth in Figure 1 at residues 4170-4179, a sequence corresponding to an upstream motif or *cis* element capable of binding zinc as set forth in Figure 1 at residues 4285-4293, a sequence corresponding to an upstream motif or *cis* element characteristic of CAP/CRP-galO as set forth in Figure 1 at residues 4379-4404, sequences corresponding to an upstream motif or *cis* element capable of binding AP1 as set forth in Figure 1 at residues 4428-4434, and 4627-4639, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding SRY as set forth in Figure 1 at residues 4625-4634, a sequence corresponding to an upstream motif or *cis* element characteristic of GC2 as set forth in Figure 1 at residues 4678-4711, a sequence corresponding to an upstream motif or *cis* element capable of binding PEA3 as set forth in Figure 1 at residues 4765-4769, a sequence corresponding to an upstream motif or *cis* element capable of MIR as set forth in Figure 1 at residues 4759-4954, a sequence corresponding to an upstream motif or *cis* element capable of binding NF-HNF-1 as set forth in Figure 1 at residues 4923-4941, a sequence corresponding to a thyroid receptor upstream motif or *cis* element as set forth in Figure 1 at residues 5151-5156, and a sequence corresponding to an upstream motif or *cis* element capable of binding NFκB as set forth in Figure 1 at residues 5166-5175 may be used to synthesize all or any portion of the TIGR promoter or any of the TIGR upstream motifs or portions the TIGR cDNA (Zamechik *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.) 83*:4143 (1986); Goodchild *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.) 85*:5507 (1988); Wickstrom *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.) 85*:1028; Holt, J.T. *et al.*,

Molec. Cell. Biol. 8:963 (1988); Gerwirtz, A.M. *et al.*, *Science 242*:1303 (1988); Anfossi, G., *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.) 86*:3379 (1989); Becker, D., *et al.*, *EMBO J. 8*:3679 (1989); all of which references are incorporated herein by reference).

Automated nucleic acid synthesizers may be employed for this purpose. In lieu of such synthesis, the disclosed SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 33, sequences corresponding to an upstream motif or *cis* element characteristic of PRL-FP111 as set forth in Figure 1 at residues 370-388, and 4491-4502, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding GR/PR as set forth in Figure 1 at residues 433-445, sequences corresponding to an upstream shear stress motif or *cis* element as set forth in Figure 1 at residues 446-451, 1288-1293, 3597-3602, 4771-4776, and 5240-5245, respectively, sequences corresponding to glucocorticoid response upstream motif or *cis* element as set forth in Figure 1 at residues 574-600, 1042-1056, 2444-2468, 2442-2269, 3536-3563, 4574-4593, 4595-4614, 4851-4865, 4844-4864, 5079-5084, 5083-5111, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding CBE as set forth in Figure 1 at residues 735-746, a sequence corresponding to an upstream motif or *cis* element capable of binding NFE as set forth in Figure 1 at residues 774-795, a sequence corresponding to an upstream motif or *cis* element capable of binding KTF.1-CS as set forth in Figure 1 at residues 843-854, a sequence corresponding to an upstream motif or *cis* element capable of binding PRE is set forth in Figure 1 at residues 987-1026, a sequence corresponding to an upstream motif or *cis* element capable of binding ETF-EGFR as set forth in Figure 1 at residues 1373-1388, a sequence corresponding to an upstream motif or *cis* element capable of binding SRE-cFos as set forth in Figure 1 at residues 1447-1456, a sequence corresponding to an upstream motif or *cis* element capable of binding Alu as set forth in Figure 1 at residues 1331-1550, a sequence corresponding to an upstream motif or *cis* element capable of binding VBP as set forth in Figure 1 at residues 1786-1797, a sequence corresponding to an upstream motif or *cis* element capable of binding Malt-CS as set forth in Figure 1 at residues 1832-1841, sequences corresponding to an upstream motif or *cis* element capable of binding ERE as set forth in Figure 1 at residues 2167-2195, 3413-3429, and 3892-3896, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding NF-mutagen as set forth in Figure 1 at residues 2329-2338, a sequence corresponding to an upstream motif or *cis* element capable of binding myc-PRF as set forth in Figure 1 at residues 2403-2416, sequences corresponding to an upstream motif or *cis* element capable of binding AP2 as set forth in Figure 1 at residues 2520-2535 and 5170-5187, respectively, sequences

corresponding to an upstream motif or *cis* element capable of binding HSTF as set forth in Figure 1 at residues 2622-2635, and 5105-5132, respectively, a sequence corresponding to an upstream motif or *cis* element characteristic of SBF as set forth in Figure 1 at residues 2733-2743, sequences corresponding to an upstream motif or *cis* element capable of binding NF-1 as set forth in Figure 1 at residues 2923-2938, 4144-4157, and 4887-4900, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding NF-MHCIIA/B as set forth in Figure 1 at residues 2936-2944, a sequence corresponding to an upstream motif or *cis* element capable of binding PEA1 as set forth in Figure 1 at residues 3285-3298, a sequence corresponding to an upstream motif or *cis* element capable of binding ICS as set forth in Figure 1 at residues 3688-3699, a sequence corresponding to an upstream motif or *cis* element capable of binding ISGF2 as set forth in Figure 1 at residues 4170-4179, a sequence corresponding to an upstream motif or *cis* element capable of binding zinc as set forth in Figure 1 at residues 4285-4293, a sequence corresponding to an upstream motif or *cis* element characteristic of CAP/CRP-galO as set forth in Figure 1 at residues 4379-4404, sequences corresponding to an upstream motif or *cis* element capable of binding AP1 as set forth in Figure 1 at residues 4428-4434, and 4627-4639, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding SRY as set forth in Figure 1 at residues 4625-4634, a sequence corresponding to an upstream motif or *cis* element characteristic of GC2 as set forth in Figure 1 at residues 4678-4711, a sequence corresponding to an upstream motif or *cis* element capable of binding PEA3 as set forth in Figure 1 at residues 4765-4769, a sequence corresponding to an upstream motif or *cis* element capable of MIR as set forth in Figure 1 at residues 4759-4954, a sequence corresponding to an upstream motif or *cis* element capable of binding NF-HNF-1 as set forth in Figure 1 at residues 4923-4941, a sequence corresponding to a thyroid receptor upstream motif or *cis* element as set forth in Figure 1 at residues 5151-5156, and a sequence corresponding to an upstream motif or *cis* element capable of binding NFκB as set forth in Figure 1 at residues 5166-5175 may be used to define a pair of primers that can be used with the polymerase chain reaction (Mullis, K. *et al.*, *Cold Spring Harbor Symp. Quant. Biol. 51*:263-273 (1986); Erlich H. *et al.*, EP 50,424; EP 84,796, EP 258,017, EP 237,362; Mullis, K., EP 201,184; Mullis K. *et al.*, US 4,683,202; Erlich, H., US 4,582,788; and Saiki, R. *et al.*, US 4,683,194)) to amplify and obtain any desired TIGR gene DNA molecule or fragment.

The TIGR promoter sequence(s) and TIGR flanking sequences can also be obtained by incubating oligonucleotide probes of TIGR oligonucleotides with members of genomic human libraries and recovering clones that hybridize to the probes. In a second embodiment, methods of "chromosome walking," or 3' or 5' RACE may be used (Frohman, M.A. *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.) 85*:8998-9002 (1988), herein incorporated by reference); Ohara, O. *et al.*, *Proc.*

Natl. Acad. Sci. (U.S.A.) 86:5673-5677 (1989), herein incorporated by reference) to obtain such sequences.

II. Uses of the Molecules of the Invention in the Diagnosis and Prognosis of Glaucoma and Related Diseases

A particularly desired use of the present invention relates to the diagnosis of glaucoma, POAG, pigmentary glaucoma, high tension glaucoma and low tension glaucoma and their related diseases. Another particularly desired use of the present invention relates to the prognosis of glaucoma, POAG, pigmentary glaucoma, high tension glaucoma and low tension glaucoma and their related diseases. As used herein the term "glaucoma" includes both primary glaucomas, secondary glaucomas, juvenile glaucomas, congenital glaucomas, and familial glaucomas, including, without limitation, pigmentary glaucoma, high tension glaucoma and low tension glaucoma and their related diseases. As indicated above, methods for diagnosing or prognosing glaucoma suffer from inaccuracy, or require multiple examinations. The molecules of the present invention may be used to define superior assays for glaucoma. Quite apart from such usage, the molecules of the present invention may be used to diagnosis or predict an individual's sensitivity to elevated intraocular pressure upon administration of steroids such as glucocorticoids or corticosteroids, or anti-inflammatory steroids). Dexamethasone, cortisol and prednisolone are preferred steroids for this purpose. Medical conditions such as inflammatory and allergic disorders, as well as organ transplantation recipients, benefit from treatment with glucocorticoids. Certain individuals exhibit an increased IOP response to such steroids (i.e., "steroid sensitivity"), which is manifested by an undesired increase in intraocular pressure. The present invention may be employed to diagnosis or predict such sensitivity, as well as glaucoma and related diseases.

In a first embodiment, the TIGR molecules of the present invention are used to determine whether an individual has a mutation affecting the level (i.e., the concentration of TIGR mRNA or protein in a sample, etc.) or pattern (i.e., the kinetics of expression, rate of decomposition, stability profile, etc.) of the TIGR expression (collectively, the "TIGR response" of a cell or bodily fluid) (for example, a mutation in the TIGR gene, or in a regulatory region(s) or other gene(s) that control or affect the expression of TIGR), and being predictive of individuals who would be predisposed to glaucoma (prognosis), related diseases, or steroid sensitivity. As used herein, the TIGR response manifested by a cell or bodily fluid is said to be "altered" if it differs from the TIGR response of cells or of bodily fluids of normal individuals. Such alteration may be manifested by either abnormally increased or abnormally diminished TIGR response. To determine whether a TIGR response is altered, the TIGR response manifested by the cell or bodily fluid of the patient is compared with that of a similar cell sample (or bodily fluid sample) of normal individuals. As will be appreciated, it is not necessary to re-determine the TIGR response

of the cell sample (or bodily fluid sample) of normal individuals each time such a comparison is made; rather, the TIGR response of a particular individual may be compared with previously obtained values of normal individuals.

In one sub-embodiment, such an analysis is conducted by determining the presence and/or identity of polymorphism(s) in the TIGR gene or its flanking regions which are associated with glaucoma, or a predisposition (prognosis) to glaucoma, related diseases, or steroid sensitivity. As used herein, the term "TIGR flanking regions" refers to those regions which are located either upstream or downstream of the TIGR coding region.

Any of a variety of molecules can be used to identify such polymorphism(s). In one embodiment, SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 33, sequences corresponding to an upstream motif or *cis* element characteristic of PRL-FP111 as set forth in Figure 1 at residues 370-388, and 4491-4502, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding GR/PR as set forth in Figure 1 at residues 433-445, sequences corresponding to an upstream shear stress motif or *cis* element as set forth in Figure 1 at residues 446-451, 1288-1293, 3597-3602, 4771-4776, and 5240-5245, respectively, sequences corresponding to glucocorticoid response upstream motif or *cis* element as set forth in Figure 1 at residues 574-600, 1042-1056, 2444-2468, 2442-2269, 3536-3563, 4574-4593, 4595-4614, 4851-4865, 4844-4864, 5079-5084, 5083-5111, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding CBE as set forth in Figure 1 at residues 735-746, a sequence corresponding to an upstream motif or *cis* element capable of binding NFE as set forth in Figure 1 at residues 774-795, a sequence corresponding to an upstream motif or *cis* element capable of binding KTF.1-CS as set forth in Figure 1 at residues 843-854, a sequence corresponding to an upstream motif or *cis* element capable of binding PRE is set forth in Figure 1 at residues 987-1026, a sequence corresponding to an upstream motif or *cis* element capable of binding ETF-EGFR as set forth in Figure 1 at residues 1373-1388, a sequence corresponding to an upstream motif or *cis* element capable of binding SRE-cFos as set forth in Figure 1 at residues 1447-1456, a sequence corresponding to an upstream motif or *cis* element capable of binding Alu as set forth in Figure 1 at residues 1331-1550, a sequence corresponding to an upstream motif or *cis* element capable of binding VBP as set forth in Figure 1 at residues 1786-1797, a sequence corresponding to an upstream motif or *cis* element capable of binding Malt-CS as set forth in Figure 1 at residues 1832-1841, sequences corresponding to an upstream motif or *cis* element capable of binding ERE as set forth in Figure 1 at residues 2167-2195, 3413-3429, and 3892-

3896, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding NF-mutagen as set forth in Figure 1 at residues 2329-2338, a sequence corresponding to an upstream motif or *cis* element capable of binding myc-PRF as set forth in Figure 1 at residues 2403-2416, sequences corresponding to an upstream motif or *cis* element capable of binding AP2 as set forth in Figure 1 at residues 2520-2535 and 5170-5187, respectively, sequences corresponding to an upstream motif or *cis* element capable of binding HSTF as set forth in Figure 1 at residues 2622-2635, and 5105-5132, respectively, a sequence corresponding to an upstream motif or *cis* element characteristic of SBF as set forth in Figure 1 at residues 2733-2743, sequences corresponding to an upstream motif or *cis* element capable of binding NF-1 as set forth in Figure 1 at residues 2923-2938, 4144-4157, and 4887-4900, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding NF-MHCIIA/B as set forth in Figure 1 at residues 2936-2944, a sequence corresponding to an upstream motif or *cis* element capable of binding PEA1 as set forth in Figure 1 at residues 3285-3298, a sequence corresponding to an upstream motif or *cis* element capable of binding ICS as set forth in Figure 1 at residues 3688-3699, a sequence corresponding to an upstream motif or *cis* element capable of binding ISGF2 as set forth in Figure 1 at residues 4170-4179, a sequence corresponding to an upstream motif or *cis* element capable of binding zinc as set forth in Figure 1 at residues 4285-4293, a sequence corresponding to an upstream motif or *cis* element characteristic of CAP/CRP-galO as set forth in Figure 1 at residues 4379-4404, sequences corresponding to an upstream motif or *cis* element capable of binding AP1 as set forth in Figure 1 at residues 4428-4434, and 4627-4639, respectively, a sequence corresponding to an upstream motif or *cis* element capable of binding SRY as set forth in Figure 1 at residues 4625-4634, a sequence corresponding to an upstream motif or *cis* element characteristic of GC2 as set forth in Figure 1 at residues 4678-4711, a sequence corresponding to an upstream motif or *cis* element capable of binding PEA3 as set forth in Figure 1 at residues 4765-4769, a sequence corresponding to an upstream motif or *cis* element capable of MIR as set forth in Figure 1 at residues 4759-4954, a sequence corresponding to an upstream motif or *cis* element capable of binding NF-HNF-1 as set forth in Figure 1 at residues 4923-4941, a sequence corresponding to a thyroid receptor upstream motif or *cis* element as set forth in Figure 1 at residues 5151-5156, and a sequence corresponding to an upstream motif or *cis* element capable of binding NFκB as set forth in Figure 1 at residues 5166-5175 (or a sub-sequence thereof) may be employed as a marker nucleic acid molecule to identify such polymorphism(s).

Alternatively, such polymorphisms can be detected through the use of a marker nucleic acid molecule or a marker protein that is genetically linked to (i.e., a polynucleotide that co-segregates with) such polymorphism(s). As stated above, the TIGR gene and/or a sequence or sequences that specifically hybridize to the TIGR gene have been mapped to chromosome 1q, 21-

32, and more preferably to the TIGR gene located at chromosome 1, q21-27, and more preferably to the TIGR gene located at chromosome 1, q22-26, and most preferably to the TIGR gene located at chromosome 1, q24. In a preferred aspect of this embodiment, such marker nucleic acid molecules will have the nucleotide sequence of a polynucleotide that is closely genetically linked to such polymorphism(s) (e.g., markers located at chromosome 1, q19-25 (and more preferably chromosome 1, q23-25, and most preferably chromosome 1, q24.

Localization studies using a Stanford G3 radiation hybrid panel mapped the TIGR gene with the D1S2536 marker nucleic acid molecules at the D1S2536 locus with a LOD score of 6.0. Other marker nucleic acid molecules in this region include: D1S210; D1S1552; D1S2536; D1S2790; SHGC-12820; and D1S2558. Other polynucleotide markers that map to such locations are known and can be employed to identify such polymorphism(s).

The genomes of animals and plants naturally undergo spontaneous mutation in the course of their continuing evolution (Gusella, J.F., *Ann. Rev. Biochem. 55*:831-854 (1986)). A "polymorphism" in the TIGR gene or its flanking regions is a variation or difference in the sequence of the TIGR gene or its flanking regions that arises in some of the members of a species. The variant sequence and the "original" sequence co-exist in the species' population. In some instances, such co-existence is in stable or quasi-stable equilibrium.

A polymorphism is thus said to be "allelic," in that, due to the existence of the polymorphism, some members of a species may have the original sequence (i.e. the original "allele") whereas other members may have the variant sequence (i.e. the variant "allele"). In the simplest case, only one variant sequence may exist, and the polymorphism is thus said to be di-allelic. In other cases, the species' population may contain multiple alleles, and the polymorphism is termed tri-allelic, etc. A single gene may have multiple different unrelated polymorphisms. For example, it may have a di-allelic polymorphism at one site, and a multi-allelic polymorphism at another site.

The variation that defines the polymorphism may range from a single nucleotide variation to the insertion or deletion of extended regions within a gene. In some cases, the DNA sequence variations are in regions of the genome that are characterized by short tandem repeats (STRs) that include tandem di- or tri-nucleotide repeated motifs of nucleotides. Polymorphisms characterized by such tandem repeats are referred to as "variable number tandem repeat" ("VNTR") polymorphisms. VNTRs have been used in identity and paternity analysis (Weber, J.L., U.S. Patent 5,075,217; Armour, J.A.L. *et al.*, *FEBS Lett. 307*:113-115 (1992); Jones, L. *et al.*, *Eur. J. Haematol. 39*:144-147 (1987); Horn, G.T. *et al.*, PCT Application WO91/14003; Jeffreys, A.J., European Patent Application 370,719; Jeffreys, A.J., U.S. Patent 5,175,082); Jeffreys. A.J. *et al.*, *Amer. J. Hum. Genet. 39*:11-24 (1986); Jeffreys. A.J. *et al.*, *Nature 316*:76-79 (1985); Gray, I.C. *et al.*, *Proc. R. Acad. Soc. Lond. 243*:241-253 (1991); Moore, S.S. *et al.*, *Genomics 10*:654-660

(1991); Jeffreys, A.J. *et al., Anim. Genet. 18*:1-15 (1987); Hillel, J. *et al., Anim. Genet. 20*:145-155 (1989); Hillel, J. *et al., Genet. 124*:783-789 (1990)).

In an alternative embodiment, such polymorphisms can be detected through the use of a marker nucleic acid molecule that is physically linked to such polymorphism(s). For this purpose, marker nucleic acid molecules comprising a nucleotide sequence of a polynucleotide located within 1 mb of the polymorphism(s), and more preferably within 100 kb of the polymorphism(s), and most preferably within 10 kb of the polymorphism(s) can be employed. Examples of such marker nucleic acids are set out in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25.

In another embodiment a marker nucleic acid will be used that is capable of specifically detecting *TIGRmt1*, *TIGRmt2*, *TIGRmt3*, *TIGRmt4*, *TIGRmt5*, *TIGRmt11*, *TIGRsv1*, or a combination of these mutations. Methods to detect base(s) substitutions, base(s) deletions and base(s) additions are known in the art (i.e. methods to genotype an individual). For example, "Genetic Bit Analysis ("GBA") method is disclosed by Goelet, P. *et al.*, WO 92/15712, herein incorporated by reference, may be used for detecting the single nucleotide polymorphisms of the present invention. GBA is a method of polymorphic site interrogation in which the nucleotide sequence information surrounding the site of variation in a target DNA sequence is used to design an oligonucleotide primer that is complementary to the region immediately adjacent to, but not including, the variable nucleotide in the target DNA. The target DNA template is selected from the biological sample and hybridized to the interrogating primer. This primer is extended by a single labeled dideoxynucleotide using DNA polymerase in the presence of two, and preferably all four chain terminating nucleoside triphosphate precursors. Cohen, D. *et al.*, (PCT Application WO91/02087) describes a related method of genotyping.

Other primer-guided nucleotide incorporation procedures for assaying polymorphic sites in DNA have been described (Komher, J. S. *et al., Nucl. Acids. Res. 17*:7779-7784 (1989), herein incorporated by reference; Sokolov, B. P., *Nucl. Acids Res. 18*:3671 (1990), herein incorporated by reference; Syvänen, A.-C., *et al., Genomics 8*:684 - 692 (1990), herein incorporated by reference; Kuppuswamy, M.N. *et al., Proc. Natl. Acad. Sci. (U.S.A.) 88*:1143-1147 (1991), herein incorporated by reference; Prezant, T.R. *et al., Hum. Mutat. 1*:159-164 (1992), herein incorporated by reference; Ugozzoli, L. *et al., GATA 9*:107-112 (1992), herein incorporated by reference; Nyrén, P. *et al., Anal. Biochem. 208*:171-175 (1993), herein incorporated by reference).

The detection of polymorphic sites in a sample of DNA may be facilitated through the use of nucleic acid amplification methods. Such methods specifically increase the concentration of polynucleotides that span the polymorphic site, or include that site and sequences located either distal or proximal to it. Such amplified molecules can be readily detected by gel electrophoresis or other means.

Another preferred method of achieving such amplification employs the polymerase chain reaction ("PCR") (Mullis, K. *et al.*, *Cold Spring Harbor Symp. Quant. Biol. 51*:263-273 (1986); Erlich H. *et al.*, European Patent Appln. 50,424; European Patent Appln. 84,796, European Patent Application 258,017, European Patent Appln. 237,362; Mullis, K., European Patent Appln. 201,184; Mullis K. *et al.*, U.S. Patent No. 4,683,202; Erlich, H., U.S. Patent No. 4,582,788; and Saiki, R. *et al.*, U.S. Patent No. 4,683,194), using primer pairs that are capable of hybridizing to the proximal sequences that define a polymorphism in its double-stranded form.

In lieu of PCR, alternative methods, such as the "Ligase Chain Reaction" ("LCR") may be used (Barany, F., *Proc. Natl. Acad. Sci. (U.S.A.) 88*:189-193 (1991). LCR uses two pairs of oligonucleotide probes to exponentially amplify a specific target. The sequences of each pair of oligonucleotides is selected to permit the pair to hybridize to abutting sequences of the same strand of the target. Such hybridization forms a substrate for a template-dependent ligase. As with PCR, the resulting products thus serve as a template in subsequent cycles and an exponential amplification of the desired sequence is obtained.

LCR can be performed with oligonucleotides having the proximal and distal sequences of the same strand of a polymorphic site. In one embodiment, either oligonucleotide will be designed to include the actual polymorphic site of the polymorphism. In such an embodiment, the reaction conditions are selected such that the oligonucleotides can be ligated together only if the target molecule either contains or lacks the specific nucleotide that is complementary to the polymorphic site present on the oligonucleotide. Alternatively, the oligonucleotides may be selected such that they do not include the polymorphic site (see, Segev, D., PCT Application WO 90/01069).

The "Oligonucleotide Ligation Assay" ("OLA") may alternatively be employed (Landegren, U. *et al.*, *Science 241*:1077-1080 (1988)). The OLA protocol uses two oligonucleotides which are designed to be capable of hybridizing to abutting sequences of a single strand of a target. OLA, like LCR, is particularly suited for the detection of point mutations. Unlike LCR, however, OLA results in "linear" rather than exponential amplification of the target sequence.

Nickerson, D.A. *et al.*, have described a nucleic acid detection assay that combines attributes of PCR and OLA (Nickerson, D.A. *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.) 87*:8923-8927 (1990). In this method, PCR is used to achieve the exponential amplification of target DNA,

which is then detected using OLA. In addition to requiring multiple, and separate, processing steps, one problem associated with such combinations is that they inherit all of the problems associated with PCR and OLA.

Schemes based on ligation of two (or more) oligonucleotides in the presence of nucleic acid having the sequence of the resulting "di-oligonucleotide", thereby amplifying the di-oligonucleotide, are also known (Wu, D.Y. *et al., Genomics 4:*560 (1989)), and may be readily adapted to the purposes of the present invention.

Other known nucleic acid amplification procedures, such as allele-specific oligomers, branched DNA technology, transcription-based amplification systems, or isothermal amplification methods may also be used to amplify and analyze such polymorphisms (Malek, L.T. *et al.,* U.S. Patent 5,130,238; Davey, C. *et al.,* European Patent Application 329,822; Schuster *et al.,* U.S. Patent 5,169,766; Miller, H.I. *et al.,* PCT appln. WO 89/06700; Kwoh, D. *et al., Proc. Natl. Acad. Sci. (U.S.A.) 86*:1173 (1989); Gingeras, T.R. *et al.,* PCT application WO 88/10315; Walker, G.T. *et al., Proc. Natl. Acad. Sci. (U.S.A.) 89*:392-396 (1992)). All the foregoing nucleic acid amplification methods could be used to predict or diagnose glaucoma.

The identification of a polymorphism in the TIGR gene, or flanking sequences up to about 5,000 base from either end of the coding region, can be determined in a variety of ways. By correlating the presence or absence of glaucoma in an individual with the presence or absence of a polymorphism in the TIGR gene or its flanking regions, it is possible to diagnose the predisposition (prognosis) of an asymptomatic patient to glaucoma, related diseases, or steroid sensitivity. If a polymorphism creates or destroys a restriction endonuclease cleavage site, or if it results in the loss or insertion of DNA (e.g., a VNTR polymorphism), it will alter the size or profile of the DNA fragments that are generated by digestion with that restriction endonuclease. As such, individuals that possess a variant sequence can be distinguished from those having the original sequence by restriction fragment analysis. Polymorphisms that can be identified in this manner are termed “restriction fragment length polymorphisms” (“RFLPs”). RFLPs have been widely used in human and animal genetic analyses (Glassberg, J., UK patent Application 2135774; Skolnick, M.H. *et al.*, Cytogen. Cell Genet. *32*:58-67 (1982); Botstein, D. *et al.*, Ann. J. Hum. Genet. *32*:314-331 (1980); Fischer, S.G *et al.* (PCT Application WO90/13668); Uhlen, M., PCT Application WO90/11369)). The role of TIGR in glaucoma pathogenesis indicates that the presence of genetic alterations (e.g., DNA polymorphisms) that affect the TIGR response can be employed to predict glaucoma .

A preferred method of achieving such identification employs the single-strand conformational polymorphism (SSCP) approach. The SSCP technique is a method capable of identifying most sequence variations in a single strand of DNA, typically between 150 and 250 nucleotides in length (Elles, Methods in Molecular Medicine: Molecular Diagnosis of Genetic

Diseases, Humana Press (1996), herein incorporated by reference); Orita *et al., Genomics 5:* 874-879 (1989), herein incorporated by reference). Under denaturing conditions a single strand of DNA will adopt a conformation that is uniquely dependent on its sequence conformation. This conformation usually will be different, even if only a single base is changed. Most conformations have been reported to alter the physical configuration or size sufficiently to be detectable by electrophoresis. A number of protocols have been described for SSCP including, but not limited to Lee *et al., Anal. Biochem. 205:* 289-293 (1992), herein incorporated by reference; Suzuki *et al., Anal. Biochem. 192:* 82-84 (1991), herein incorporated by reference; Lo *et al., Nucleic Acids Research 20:* 1005-1009 (1992), herein incorporated by reference; Sarkar *et al., Genomics 13:* 441-443 (1992), herein incorporated by reference).

In accordance with this embodiment of the invention, a sample DNA is obtained from a patient. In a preferred embodiment, the DNA sample is obtained from the patient's blood. However, any source of DNA may be used. The DNA is subjected to restriction endonuclease digestion. TIGR is used as a probe in accordance with the above-described RFLP methods. By comparing the RFLP pattern of the TIGR gene obtained from normal and glaucomatous patients, one can determine a patient's predisposition (prognosis) to glaucoma. The polymorphism obtained in this approach can then be cloned to identify the mutation at the coding region which alters the protein's structure or regulatory region of the gene which affects its expression level. Changes involving promoter interactions with other regulatory proteins can be identified by, for example, gel shift assays using HTM cell extracts, fluid from the anterior chamber of the eye, serum, etc. Interactions of TIGR protein in glaucomatous cell extracts, fluid from the anterior chamber of the eye, serum, etc. can be compared to control samples to thereby identify changes in those properties of TIGR that relate to the pathogenesis of glaucoma. Similarly such extracts and fluids as well as others (blood, etc.) can be used to diagnosis or predict steroid sensitivity.

Several different classes of polymorphisms may be identified through such methods. Examples of such classes include: (1) polymorphisms present in the TIGR cDNA of different individuals; (2) polymorphisms in non-translated TIGR gene sequences, including the promoter or other regulatory regions of the TIGR gene; (3) polymorphisms in genes whose products interact with TIGR regulatory sequences; (4) polymorphisms in gene sequences whose products interact with the TIGR protein, or to which the TIGR protein binds.

In an alternate sub-embodiment, the evaluation is conducted using oligonucleotide "probes" whose sequence is complementary to that of a portion of SEQ ID NO: 1, SEQ ID NO: 2 SEQ ID NO: 3, SEQ ID NO: 4, or SEQ ID NO: 5. Such molecules are then incubated with cell extracts of a patient under conditions sufficient to permit nucleic acid hybridization.

In one sub-embodiment of this aspect of the present invention, one can diagnose or predict glaucoma, related diseases and steroid sensitivity by ascertaining the TIGR response in a biopsy

(or a macrophage or other blood cell sample), or other cell sample, or more preferably, in a sample of bodily fluid (especially, blood, serum, plasma, tears, buccal cavity, etc.). Since the TIGR gene is induced in response to the presence of glucocorticoids, a highly preferred embodiment of this method comprises ascertaining such TIGR response prior to, during and/or subsequent to, the administration of a glucocorticoid. Thus, by way of illustration, glaucoma could be diagnosed or predicted by determining whether the administration of a glucocorticoid (administered topically, intraocularly, intramuscularly, systemically, or otherwise) alters the TIGR response of a particular individual, relative to that of normal individuals. Most preferably, for this purpose, at least a "TIGR gene-inducing amount" of the glucocorticoid will be provided. As used herein, a TIGR gene-inducing amount of a glucocorticoid is an amount of glucocorticoid sufficient to cause a detectable induction of TIGR expression in cells of glaucomatous or non-glaucomatous individuals.

Generating Cells, Vectors, and Expressed Proteins Using Agents of the Invention

The present invention also relates to methods for obtaining a recombinant host cell, especially a mammalian host cell, comprising introducing into a host cell exogenous genetic material comprising a nucleic acid of the invention. The present invention also relates to an insect cell comprising a recombinant vector having a nucleic acid of the invention. The present invention also relates to methods for obtaining a recombinant host cell, comprising introducing exogenous genetic material comprising a nucleic acid of the invention via homologous recombination. Through homologous recombination, the promoter and 5' flanking sequences of the TIGR gene described here can be used in gene activation methods to produce a desired gene product in host cells (*see, for example,* U.S. Patent 5,733,746, specifically incorporated herein by reference). The specific expression of the TIGR gene in TM cells afforded by the TIGR promoter region DNA can, thus, be transferred via homologous recombination to express other gene products in a similar fashion. Some of these other gene products may be therapeutic proteins that address diseases related to increased IOP or glaucoma. Methods for selecting and using the promoter and 5' flanking sequence for the gene targeting technique involved in the gene activation method are known in the art. Depending upon the nature of the modification and associated targeting construct, various techniques may be employed for identifying targeted integration. Conveniently, the DNA may be digested with one or more restriction enzymes and the fragments probed with an appropriate DNA fragment, which will identify the properly sized restriction fragment associated with integration.

The sequence to be integrated into the host may be introduced by any convenient means, which includes calcium precipitated DNA, spheroplast fusion, transformation, electroporation, biolistics, lipofection, microinjection, or other convenient means. Where an amplifiable gene is being employed, the amplifiable gene may serve as the selection marker for selecting hosts into which the amplifiable gene has been introduced. Alternatively, one may include with the amplifiable gene another marker, such as a drug resistance marker, e.g. neomycin resistance (G418 in mammalian cells), hygromycin resistance etc., or an auxotrophy marker (HIS3, TRP1, LEU2, URA3, ADE2, LYS2, etc.) for use in yeast cells.

For example, homologous recombination constructs can be prepared where the amplifiable gene will be flanked, normally on both sides, with DNA homologous with the DNA of the target region, here the TIGR sequences. Depending upon the nature of the integrating DNA and the purpose of the integration, the homologous DNA will generally be within 100 kb, usually 50 kb, preferably about 25 kb, of the transcribed region of the target gene, more preferably within 2 kb of the target gene. The homologous DNA may include the 5'-upstream region outside of the transcriptional regulatory region or enhancer sequences, transcriptional initiation sequences, adjacent sequences, or the like. The homologous region may include a portion of the coding region, where the coding region may be comprised only of an open reading frame or of combination of exons and introns. The homologous region may also comprise all or a portion of an intron, where all or a portion of one or more exons may also be present. Alternatively, the homologous region may comprise the 3'-region, so as to comprise all or a portion of the transcriptional termination region, or the region 3' of this region. The homologous regions may extend over all or a portion of the target gene or be outside the target gene comprising all or a portion of the transcriptional regulatory regions and/or the structural gene.

The integrating constructs may be prepared in accordance with conventional ways, where sequences may be synthesized, isolated from natural sources, manipulated, cloned, ligated, subjected to in vitro mutagenesis, primer repair, or the like. At various stages, the joined sequences may be cloned, and analyzed by restriction analysis, sequencing, or the like. Usually during the preparation of a construct where various fragments are joined, the fragments, intermediate constructs and constructs will be carried on a cloning vector comprising a replication system functional in a prokaryotic host, e.g., *E. coli*, and a marker for selection, e.g., biocide resistance, complementation to an auxotrophic host, etc. Other functional sequences may also be present, such as polylinkers, for ease of introduction and excision of the construct or portions thereof, or the like. A large number of cloning vectors are available such as pBR322, the pUC series, etc. These constructs may then be used for integration into the primary host.

DNA comprising a nucleic acid of the invention can be introduced into a host cell by a variety of techniques that include calcium phosphate/DNA co-precipitates, microinjection of DNA

into the nucleus, electroporation, yeast protoplast fusion with intact cells, transfection, polycations, e.g., polybrene, polyornithine, etc., or the like. The DNA may be single or double stranded DNA, linear or circular. The various techniques for transforming cells are well known (see Keown *et al., Methods Enzymol.* (1989), Keown *et al., Methods Enzymol. 185*:527-537 (1990); Mansour *et al., Nature 336*:348-352, (1988); all of which are herein incorporated by reference in their entirety).

In a preferred aspect, the invention relates to recombinant insect vectors and insect cells comprising a nucleic acid of the invention. In a particularly preferred aspect, a Baculovirus expression vector is used, introduced into an insect cell, and recombinant TIGR protein expressed. The recombinant TIGR protein may be the full length protein from human TM endothelial cells, a fusion protein comprising a substantial fragment of the full length protein, for example, at least about 20 contiguous amino acids to about 100 contiguous amino acids of the full length protein, or a variant TIGR protein or fusion protein produced by site-directed mutagenesis, DNA shuffling, or a similar technique. Generally, the variant TIGR proteins and the fusion proteins will retain at least one structural or functional characteristic of the full length TIGR protein, such as the ability to bind the same antibody, the presence of the substantially similar leucine zipper region, or the ability to bind the same ligand or receptor on TM cells (*see* Nguyen *et al.*, *J. Biol. Chem.* 273:6341-6350 (1998), specifically incorporated herein by reference). Nucleic acids comprising the leucine zipper-encoding regions of the TIGR gene can be identified by methods known in the art and can be used in combination with recombinant or synthetic methods to create ligand-receptor assays.

Examples of the preferred, recombinant insect vector, host cell, and TIGR protein of the invention were generated by ligating TIGR cDNA into the PVL1393 vector [Invitrogen]. This vector was transferred into Sf9 cells, the TIGR protein expressed and then purified (see U.S. Patent 5,789,169 and Nguyen *et al.*, *J. Biol. Chem.* 273:6341-6350 (1998), both of which are specifically incorporated herein by reference in their entirety). An SDS-PAGE gel of the resulting proteins showed protein bands in the 55 kDa range, which were sequenced to confirm correct identity.

In preferred embodiments of the vectors, cells and related methods of the invention, a TIGR fusion protein with GFP (green fluorescent protein) can be expressed in a TM cell line (*see* Nguyen, *et al.*, *J. Biol. Chem.* 273:6341-6350 (1998) and the references cited therein for primary TM cell culture and transfection methods). Transformed, cultured TM cells at log phase were transfected with a TIGR-GFP fusion protein-encoding vector. The vector includes the CMV promoter to allow high expression, TIGR cDNA from the first ATG to the end of the protein-encoding region, a fluorescent protein tag (GFP) fused to the carboxy terminus of the TIGR-encoding sequence, and the G418 resistance gene. These elements, and their use, is known in the

art or provided by this disclosure and its incorporated references. The construct is termed TIGR1-GFP. The transfection was performed using calcium phosphate or Lipofectin techniques, as known in the art. Incubation at growth condition of 37°C, 8% CO_2, for 6-18 hours followed. After the transfection, the DNA media was replaced by fresh growth media including G418, which was changed twice weekly, until resistant colonies of cells outgrew the monolayer cells (about 10-15 days). The cell colonies were collected and propagated several passes to select for resistant, transformed cells. The expression of fluorescent TIGR-GFP fusion protein was tested for after several passes. One out of twenty selected colonies expressed high levels of the TIGR-GFP fusion protein.

In other preferred embodiments of the cells and methods of the invention, a transformed, immortalized TM cell line can be prepared using an SV40-derived vector. Primary cultured TM cells are transfected with an SV40 vector with a defect in the PsvOri, as known in the art. Briefly, primary cultured cells at log phase are transfected with PsvOri DNA using calcium phosphate or Lipofectin and incubated at growth condition of 37°C, 8% $C0_2$ for 6-18 hours. The DNA media was replaced by fresh growth media and changed twice weekly until colonies of immortalized cells outgrow the dying monolayer (about 10-15 days). The cell colonies are collected and propagated several passes to select for transformed cells.

III. Methods of Administration

Some of the agents of the present invention can be formulated according to known methods to prepare pharmacologically acceptable compositions, whereby these materials, or their functional derivatives, having the desired degree of purity are combined in admixture with a physiologically acceptable carrier, excipient, or stabilizer. Such materials are non-toxic to recipients at the dosages and concentrations employed. The active component of such compositions may be agents, analogs or mimetics of such molecules. Where nucleic acid molecules are employed, such molecules may be sense, antisense or triplex oligonucleotides of the TIGR promoter, TIGR cDNA, TIGR intron, TIGR exon or TIGR gene.

A composition is said to be "pharmacologically acceptable" if its administration can be tolerated by a recipient patient. An agent is physiologically significant if its presence results in a detectable change in the physiology of a recipient patient.

Suitable vehicles and their formulation, inclusive of other human proteins, e.g., human serum albumin, are described, for example, in Remington's Pharmaceutical Sciences (16th ed., Osol, A., Ed., Mack, Easton PA (1980)).

If the composition is to be water soluble, it may be formulated in a buffer such as phosphate or other organic acid salt preferably at a pH of about 7 to 8. If the composition is only partially soluble in water, it may be prepared as a microemulsion by formulating it with a nonionic surfactant such as Tween, Pluronics, or PEG, e.g., Tween 80, in an amount of, for example, 0.04-0.05% (w/v), to increase its solubility. The term "water soluble" as applied to the polysaccharides and polyethylene glycols is meant to include colloidal solutions and dispersions. In general, the solubility of the cellulose derivatives is determined by the degree of substitution of ether groups, and the stabilizing derivatives useful herein should have a sufficient quantity of such ether groups per anhydroglucose unit in the cellulose chain to render the derivatives water soluble. A degree of ether substitution of at least 0.35 ether groups per anhydroglucose unit is generally sufficient. Additionally, the cellulose derivatives may be in the form of alkali metal salts, for example, the Li, Na, K or Cs salts.

Optionally other ingredients may be added such as antioxidants, e.g., ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinyl pyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, mannose, or dextrins; chelating agents such as EDTA; and sugar alcohols such as mannitol or sorbitol.

Additional pharmaceutical methods may be employed to control the duration of action. Controlled or sustained release preparations may be achieved through the use of polymers to complex or absorb the TIGR molecule(s) of the composition. The controlled delivery may be exercised by selecting appropriate macromolecules (for example polyesters, polyamino acids, polyvinyl pyrrolidone, ethylenevinylacetate, methylcellulose, carboxymethylcellulose, or protamine sulfate) and the concentration of macromolecules as well as the methods of incorporation in order to control release.

Sustained release formulations may also be prepared, and include the formation of microcapsular particles and implantable articles. For preparing sustained-release compositions, the TIGR molecule(s) of the composition is preferably incorporated into a biodegradable matrix or microcapsule. A suitable material for this purpose is a polylactide, although other polymers of poly-(a-hydroxycarboxylic acids), such as poly-D-(-)-3-hydroxybutyric acid (EP 133,988A), can be used. Other biodegradable polymers include poly(lactones), poly(orthoesters), polyamino acids, hydrogels, or poly(orthocarbonates) poly(acetals). The polymeric material may also comprise polyesters, poly(lactic acid) or ethylene vinylacetate copolymers. For examples of sustained release compositions, see U.S. Patent No. 3,773,919, EP 58,481A, U.S. Patent No.

3,887,699, EP 158,277A, Canadian Patent No. 1176565, Sidman, U. *et al.*, *Biopolymers* 22:547 (1983), and Langer, R. *et al.*, *Chem. Tech.* 12:98 (1982).

Alternatively, instead of incorporating the TIGR molecule(s) of the composition into polymeric particles, it is possible to entrap these materials in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatine-microcapsules and poly(methylmethacylate) microcapsules, respectively, or in colloidal drug delivery systems, for example, liposomes, albumin microspheres, microemulsions, nanoparticles, and nanocapsules or in macroemulsions. Such techniques are disclosed in Remington's Pharmaceutical Sciences (1980).

In an alternative embodiment, liposome formulations and methods that permit intracellular uptake of the molecule will be employed. Suitable methods are known in the art, see, for example, Chicz, R.M. *et al.* (PCT Application WO 94/04557), Jaysena, S.D. *et al.* (PCT Application WO93/12234), Yarosh, D.B. (U.S. Patent No. 5,190,762), Callahan, M.V. *et al.* (U.S. Patent No. 5,270,052) and Gonzalezro, R.J. (PCT Application 91/05771), all herein incorporated by reference.

Having now generally described the invention, the same will be more readily understood through reference to the following examples which are provided by way of illustration, and are not intended to be limiting of the present invention, unless specified.

EXAMPLE 1

Illustrative Single Strand Conformational Polymorphism Assay

Single strand conformational polymorphism (SSCP) screening is carried out according to the procedure of Hue *et al., The Journal of Investigative Ophthalmology 105.4:* 529-632 (1995), herein incorporated by reference. SSCP primers are constructed corresponding to sequences found within the TIGR promoter and two of exons of TIGR. The following primers are constructed: forward primer "Sk-1a": 5'-TGA GGC TTC CTC TGG AAA C-3' (SEQ ID NO: 6); reverse primer "ca2": 5'-TGA AAT CAG CAC ACC AGT AG-3' (SEQ ID NO: 7); forward primer "CA2": 5'-GCA CCC ATA CCC CAA TAA TAG-3' (SEQ ID NO: 8); reverse primer "Pr+1": 5'-AGA GTT CCC CAG ATT TCA CC-3' (SEQ ID NO: 9); forward primer "Pr-1": 5'-ATC TGG GGA ACT CTT CTC AG-3' (SEQ ID NO: 10); reverse primer "Pr+2(4A2)": 5'-TAC AGT TGT TGC AGA TAC G-3' (SEQ ID NO: 11); forward primer "Pr-2(4A)": 5'-ACA ACG TAT CTG CAA CAA CTG-3' (SEQ ID NO: 12); reverse primer "Pr+3(4A)": 5'-TCA GGC TTA ACT GCA GAA CC-3' (SEQ ID NO: 13); forward primer "Pr-3(4A)": 5'-TTG GTT CTG CAG TTA AGC C-3' (SEQ ID NO: 14); reverse primer "Pr+2(4A1)": 5'-AGC AGC ACA AGG GCA ATC C-3' (SEQ ID NO: 15); reverse primer "Pr+1(4A)": 5'-ACA GGG CTA TAT TGT

GGG-3' (SEQ ID NO: 16); forward primer "KS1X": 5'-CCT GAG ATG CCA GCT GTC C-3' (SEQ ID NO: 17); reverse primer "SK1XX": 5'-CTG AAG CAT TAG AAG CCA AC-3' (SEQ ID NO: 18); forward primer "KS2a1": 5'-ACC TTG GAC CAG GCT GCC AG-3' (SEQ ID NO: 19); reverse primer "SK3" 5'-AGG TTT GTT CGA GTT CCA G-3' (SEQ ID NO: 20); forward primer "KS4": 5'-ACA ATT ACT GGC AAG TAT GG-3' (SEQ ID NO: 21); reverse primer "SK6A": 5'-CCT TCT CAG CCT TGC TAC C-3' (SEQ ID NO: 22); forward primer "KS5": 5'-ACA CCT CAG CAG ATG CTA CC-3' (SEQ ID NO: 23); reverse primer "SK8": 5'-ATG GAT GAC TGA CAT GGC C-3' (SEQ ID NO: 24); forward primer "KS6": 5'-AAG GAT GAA CAT GGT CAC C-3' (SEQ ID NO: 25).

The locations of primers: Sk-1a, ca2, CA2, Pr+1, Pr-1, Pr+2(4A2), Pr-2(4A), Pr+3(4A), Pr-3 (4A), Pr-3(4A), Pr+2(4A1), and Pr+1(4A) are diagramatically set forth in Figure 4. The location of primers: KS1X, SK1XX, Ks2a1, SK3, KS4, SK6A, KS5, SK8, and KS6 are diagramatically set forth in Figure 5.

Families with a history of POAG in Klamath Falls, Oregon, are screened by SSCP according to the method of Hue *et al., The Journal of Investigative Ophthalmology 105.4:* 529-632 (1995), herein incorporated by reference). SSCP primers SK-1a, ca2, CA2, Pr+1, Pr-2(4A), Pr+3(4A), SK1XX, and KS6 detect single strand conformational polymorphisms in this population. An SSCP is detected using SSCP primers Pr+3(4A) and Pr-2(4A). 70 family members of the Klamath Fall, Oregon are screened with these primers and the results are set forth in Table 1.

TABLE 1

	Total	SSCP+	SSCP-
Glaucoma positive individuals[1]	12	12	0
Glaucoma negative individuals	13	0	13
Spouses (glaucoma negative)	16	2	14
Others [2]	29	6	23

1 = glaucoma positive individuals as determined by IOP of greater than 25 mmHg

2 = unidentified glaucoma due to the age of the individual.

A second SSCP is detected using SSCP primers Pr+1 and CA2. 14 family members of the Klamath Fall, Oregon are screened with these primers. A characteristic polymorphism is found in the 6 affected family members but absent in the 8 unaffected members. A third SSCP is detected using SSCP primers ca2 and sk-1a. The same 14 family members of the Klamath Fall, Oregon that are screened with Pr+1 and CA2 are screened with ca2 and sk-1a primers. A characteristic polymorphism is found in the 6 affected family members but absent in the 8 unaffected members.

A fourth SSCP is detected using SSCP primers KS6 and SK1XX. 22 family members of the Klamath Fall, Oregon and 10 members of a Portland, Oregon pedigree are screened with these primers. A polymorphism is found in exon 3. The results are as set forth in Table 2.

TABLE 2

	Total	SSCP+	SSCP-
Klamath Fall, Oregon			
Glaucoma positive individuals[1]	3	3	0
Glaucoma negative individuals	6	0	6
Others [2]	13	6	7
Portland, Oregon			
Glaucoma positive individuals[1]	6	6	0
Glaucoma negative individuals	4	0	4
Others [2]	0	0	0

1 = glaucoma positive individuals as determined by IOP of greater than 25 mmHg

2 = unidentified glaucoma due to the age of the individual.

EXAMPLE 2

TIGR Homologies

A novel "myosin-like" acidic protein termed myocilin is expressed predominantly in the photoreceptor cells of retina and is localized particularly in the rootlet and basal body of connecting cilium (Kubota *et al.*, Genomics 41: 360-369 (1997), herein incorporated by reference). The myocilin gene is mapped to human chromosome Iq23-q24. The coding region of myocilin is 100 percent homologous with TIGR.

Homology searches are performed by GCG (Genetics Computer Group, Madison, WI) and include the GenBank, EMBL, Swiss-Prot databases and EST analysis. Using the Blast search, the best fits are found with a stretch of 177 amino acids in the carboxy terminals for an extracellular mucus protein of the olfactory, olfactomedin and three olfactomedin-like species. The alignment presented in Figure 6 shows the TIGR homology (SEQ ID NO. 27) to an expressed sequence tag (EST) sequence from human brain (ym08h12.r1)(SEQ ID NO. 28)(The WashU-Merck EST Project, 1995); the Z domain of olfactomedin-related glycoprotein from rat brain (1B426bAMZ)(SEQ ID NO. 29)(Danielson *et al.*, *Journal of Neuroscience Research* 38: 468-478 (1994), herein incorporated by reference) and the olfactomedin from olfactory tissue of bullfrogs (ranofm) (SEQ ID NO. 30)(Yokoe and Anholt, *Proc. Natl. Acad. Sci. 90*: 4655-4659 (1993), herein incorporated by reference; Snyder and Anholt, *Biochemistry 30*: 9143-9153 (1991), herein incorporated by reference). These domains share very similar amino acid positions as depicted in the consensus homology of Figure 6 (SEQ ID NO. 31), with the exception being the truncated human clone in which the position with respect to its full length sequence has not been established. No significant homology is found for the amino termini of these molecules.

EXAMPLE 3

Identification of TIGRmt11

DNA samples were obtained from individuals noted for having elevated IOP in response to the administration of topical corticosteroids. Typically, the "Armaly" criteria is used to register IOP changes.

Genomic DNA from blood or buccal swabs were used for PCR amplification. The PCR reaction includes 95° C for 30 sec, for denaturation, 55° C for30 sec, for annealing and 72° C for 30 sec for synthosis. The reaction was performed for 30 cycles with an additional cycle of 72° C for 5 min at the end.

The primer pair for the PCR reaction can include any pair that amplifies a specific region targeted for analyzing mutants or polymorphisms. Preferably, the amplified region will be from about 500 base pairs 5' of the start of transcription to the start of translation. More preferably, it will include an amplified region about 200 bp 5' of the start of transcription to about 10 base pairs 5' to the start of translation. Methods for determining amplification primer sequences from within a known sequence region are well known in the art. Examplary methods include, but are not limited to, computer generated searches using programs such as Primer3 (www-genome.wi.mit.edu/cgi-bin/primer/primer3.cgi), STSPipeline (www-genome.wi.mit.edu/cgi-bin/www-STS_Pipeline), or GeneUp (Pesole, *et al.*, *BioTechniques* *25*:112-123 (1998)).

In an especially preferred embodiment, this amplified region will be from position 5044 of SEQ ID NO: 3 to about 5327 of SEQ ID NO: 3, which will thus employ primers of the sequence of about 5044 to about 5066 and the sequence of about 5309 to about 5327 of SEQ ID NO: 3, or the complement. In one embodiment, the complement of the sequence from about 5309 to about 5327 is used as one of the primers and the sequence from about 5044 to about 5066 is used as the other primer.

For this example, the following primers were used: forward primer CA-2R (SEQ ID NO: 35 – 5' AACTATTATT GGGGTATGGG) and reverse primer Sk-la (SEQ ID NO: 36 - 5' TTGGTGAGGC TTCCTCTGC). The primers were labeled with a fluorescent dye IRD-800 by Li-Cor Technology and the PCR product (about 300 bp) was denatured by heat and subject to BESS assays to detect mutations.

BESS, or Base Excision Sequence Scanning, employed specific restriction enzyme that cleaves T position of single strand DNA. The cleavage will produce DNA fragments that could be observed by acrylamide gels. Based on this, a 'T mutation' will produce different cleavage pattern for the mutated strand compared to the normal strand. Since 95% of mutations involve a T mutation, this method is very practical. In addition to BESS, the amplified fragments can also be sequenced or compared by hybridization methods (microarray hybridization techniques or the sequencing-by-hybridization technique) in order to determine the exact nucleotide sequence, as known in the art.

Using this assay, patients exhibiting an increased IOP in response to topical corticosteroid treatments had an elevated level of a T mutation in one particular position, at about 160 bases 5' to the start of the TIGR coding region. The presence of this particular mutation, called TIGRmt11, therefore, indicated a specific genetic linkage to steroid sentivity that manifests in atleast a higher risk of increased IOP, and thus glaucoma, in repsonse to steroid treatment.

TABLE 3

Subject	Duration of CS Treatment	IOP (OD/OS)	Genotype (mt.11)
1	1 year	38/30	+/-
2	3 weeks	25/28	+/+
3	2 weeks	28/28	+/+

CS= corticosteroid, topical treatment

(1 year) CS treatment 38/30 mm Hg, OD/OS; (3 weeks) CS treatment 25/28 mm Hg, OD/OS; (2 weeks) CS treatment 28/38 mm Hg, OD/OS

The sequence in SEQ ID NO: 33 (**CAAACAGACT TCCGGAAGGT**) identifies bases immediately adjacent to the single base polymorphism, which represents bases 5101 to 5120 of SEQ ID NO: 1, except that the underlined C in the TIGRmt11 sequence variant is substituted for the 'wild type' T, found in SEQ ID NO: 1.

EXAMPLE 4

Verification of Linkage Between TIGRmt11 and Risk of Glaucoma

Subjects are given standard topical dexamethasone eye drops (0.1%) four times a day, for four weeks. Pre-treatment and post-treatment IOP readings are taken and patients are classified as having high (>16mmHg), intermediate (6-16mmHg) or low (<6mmHg) IOP responses under the "Armaly" criteria. DNA samples are obtained from four subjects having high or intermediate IOP changes. Samples from several non-responder patients were also taken. The DNA samples were analyzed for the presence of the TIGRmt11 variant sequence, as discussed above. The results are given in Table 4.

TABLE 4

Subject	Age	Classification	CS-IOP Response	Genotype (mt.11)
1	47	OHT	Intermediate	+/+
2	28	POAG	High	+/+
3	46	POAG/OHT	High	+/+
4	15	Stevens-Johnson	High	+/+
5	Nr	Normal	Low	-/-

6	Nr	Normal	Low	-/-
7	Nr	Normal	Low	-/-

OHT = Ocular Hypertensive (began with a mild IOP elevation, no POAG)
POAG = Original diagnosis is primary open-angle glaucoma
POAG/OHT = Converted to POAG, from original diagnosis OHT

The data obtained indicates the association of TIGRmt.11 and the response to topical CS. Clearly, all the subjects with clinically identifiable responses to the CS treatment possessed the TIGRmt11 variant sequence while none of the subjects with the 'wild type'sequence, or a sequence that did not possess the TIGRmt11 variant, did not.

While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications and this application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure as come within known or customary practice within the art to which the invention pertains and as may be applied to the essential features herein before set forth and as follows in the scope of the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Nguyen, Thai D.
Polansky, Jon R.
Chen, Pu
Chen, Hua

(ii) TITLE OF THE INVENTION: NUCLEIC ACIDS, KITS, AND METHODS FOR THE DIAGNOSIS, PROGNOSIS AND TREATMENT OF GLAUCOMA AND RELATED DISORDERS

(iii) NUMBER OF SEQUENCES: 36

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Howrey & Simon
(B) STREET: 1299 Pennsylvania Avenue, N.W.
(C) CITY: Washington
(D) STATE: DC
(E) COUNTRY: USA
(F) ZIP: 20004-2402

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/791,154
(B) FILING DATE: 28-JAN-1997

Sub a1

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME:
(B) REGISTRATION NUMBER:
(C) REFERENCE/DOCKET NUMBER: 07425-0051

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 202 783-0800
(B) TELEFAX: 202 383-6610
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5300 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATCTTTGTTC AGTTTACCTC AGGGCTATTA TGAAATGAAA TGAGATAACC AATGTGAAAG 60

TCCTATAAAC TGTATAGCCT CCATTCGGAT GTATGTCTTT GGCAGGATGA TAAAGAATCA 120
GGAAGAAGGA GTATCCACGT TAGCCAAGTG TCCAGGCTGT GTCTGCTCTT ATTTTAGTGA 180
CAGATGTTGC TCCTGACAGA AGCTATTCTT CAGGAAACAT CACATCCAAT ATGGTAAATC 240
CATCAAACAG GAGCTAAGAA ACAGGAATGA GATGGGCACT TGCCCAAGGA AAAATGCCAG 300
GAGAGCAAAT AATGATGAAA AATAAACTTT TCCCTTTGTT TTTAATTTCA GGAAAAAATG 360
ATGAGGACCA AAATCAATGA ATAAGGAAAA CAGCTCAGAA AAAAGATGTT TCCAAATTGG 420
TAATTAAGTA TTTGTTCCTT GGGAAGAGAC CTCCATGTGA GCTTGATGGG AAAATGGGAA 480
AAACGTCAAA AGCATGATCT GATCAGATCC CAAAGTGGAT TATTATTTTA AAAACCAGAT 540
GGCATCACTC TGGGGAGGCA AGTTCAGGAA GGTCATGTTA GCAAAGGACA TAACAATAAC 600
AGCAAAATCA AAATTCCGCA AATGCAGGAG GAAAATGGGG ACTGGGAAAG CTTTCATAAC 660
AGTGATTAGG CAGTTGACCA TGTTCGCAAC ACCTCCCCGT CTATACCAGG GAACACAAAA 720
ATTGACTGGG CTAAGCCTGG ACTTTCAAGG GAAATATGAA AAACTGAGAG CAAAACAAAA 780
GACATGGTTA AAAGGCAACC AGAACATTGT GAGCCTTCAA AGCAGCAGTG CCCCTCAGCA 840
GGGACCCTGA GGCATTTGCC TTTAGGAAGG CCAGTTTTCT TAAGGAATCT TAAGAAACTC 900
TTGAAAGATC ATGAATTTTA ACCATTTTAA GTATAAAACA AATATGCGAT GCATAATCAG 960
TTTAGACATG GGTCCCAATT TTATAAAGTC AGGCATACAA GGATAACGTG TCCCAGCTCC 1020
GGATAGGTCA GAAATCATTA GAAATCACTG TGTCCCCATC CTAACTTTTT CAGAATGATC 1080
TGTCATAGCC CTCACACACA GGCCCGATGT GTCTGACCTA CAACCACATC TACAACCCAA 1140
GTGCCTCAAC CATTGTTAAC GTGTCATCTC AGTAGGTCCC ATTACAAATG CCACCTCCCC 1200
TGTGCAGCCC ATCCCGCTCC ACAGGAAGTC TCCCCACTCT AGACTTCTGC ATCACGATGT 1260
TACAGCCAGA AGCTCCGTGA GGGTGAGGGT CTGTGTCTTA CACCTACCTG TATGCTCTAC 1320
ACCTGAGCTC ACTGCAACCT CTGCCTCCCA GGTTCAAGCA ATTCTCCTGT CTCAGCCTCC 1380
CGCGTAGCTG GGACTACAGG CGCACGCCCG GCTAATTTTT GTATTGTTAG TAGAGATGGG 1440
GTTTCACCAT ATTAGCCCGG CTGGTCTTGA ACTCCTGACC TCAGGTGATC CACCCACCTC 1500
AGCCTCCTAA AGTGCTGGGA TTACAGGCAT GAGTCACCGC GCCCGGCCAA GGGTCAGTGT 1560
TTAATAAGGA ATAACTTGAA TGGTTTACTA AACCAACAGG GAAACAGACA AAAGCTGTGA 1620
TAATTTCAGG GATTCTTGGG ATGGGGAATG GTGCCATGAG CTGCCTGCCT AGTCCCAGAC 1680
CACTGGTCCT CATCACTTTC TTCCCTCATC CTCATTTTCA GGCTAAGTTA CCATTTTATT 1740
CACCATGCTT TTGTGGTAAG CCTCCACATC GTTACTGAAA TAAGAGTATA CATAAACTAG 1800
TTCCATTTGG GGCCATCTGT GTGTGTGTAT AGGGGAGGAG GGCATACCCC AGAGACTCCT 1860
TGAAGCCCCC GGCAGAGGTT TCCTCTCCAG CTGGGGGAGC CCTGCAAGCA CCCGGGGTCC 1920
TGGGTGTCCT GAGCAACCTG CCAGCCCGTG CCACTGGTTG TTTTGTTATC ACTCTCTAGG 1980
GACCTGTTGC TTTCTATTTC TGTGTGACTC GTTCATTCAT CCAGGCATTC ATTGACAATT 2040
TATTGAGTAC TTATATCTGC CAGACACCAG AGACAAAATG GTGAGCAAAG CAGTCACTGC 2100
CCTACCTTCG TGGAGGTGAC AGTTTCTCAT GGAAGACGTG CAGAAGAAAA TTAATAGCCA 2160
GCCAACTTAA ACCCAGTGCT GAAAGAAAGG AAATAAACAC CATCTTGAAG AATTGTGCGC 2220
AGCATCCCTT AACAAGGCCA CCTCCCTAGC GCCCCCTGCT GCCTCCATCG TGCCCGGAGG 2280
CCCCCAAGCC CGAGTCTTCC AAGCCTCCTC CTCCATCAGT CACAGCGCTG CAGCTGGCCT 2340
GCCTCGCTTC CCGTGAATCG TCCTGGTGCA TCTGAGCTGG AGACTCCTTG GCTCCAGGCT 2400
CCAGAAAGGA AATGGAGAGG GAAACTAGTC TAACGGAGAA TCTGGAGGGG ACAGTGTTTC 2460
CTCAGAGGGA AAGGGGCCTC CACGTCCAGG AGAATTCCAG GAGGTGGGGA CTGCAGGGAG 2520
TGGGGACGCT GGGGCTGAGC GGGTGCTGAA AGGCAGGAAG GTGAAAAGGG CAAGGCTGAA 2580
GCTGCCCAGA TGTTCAGTGT TGTTCACGGG GCTGGGAGTT TTCCGTTGCT TCCTGTGAGC 2640
Sub a1 cont
CTTTTTATCT TTTCTCTGCT TGGAGGAGAA GAAGTCTATT TCATGAAGGG ATGCAGTTTC 2700
ATAAAGTCAG CTGTTAAAAT TCCAGGGTGT GCATGGGTTT TCCTTCACGA AGGCCTTTAT 2760
TTAATGGGAA TATAGGAAGC GAGCTCATTT CCTAGGCCGT TAATTCACGG AAGAAGTGAC 2820
TGGAGTCTTT TCTTTCATGT CTTCTGGGCA ACTACTCAGC CCTGTGGTGG ACTTGGCTTA 2880
TGCAAGACGG TCGAAAACCT TGGAATCAGG AGACTCGGTT TTCTTTCTGG TTCTGCCATT 2940
GGTTGGCTGT GCGACCGTGG GCAAGTGTCT CTCCTTCCCT GGGCCATAGT CTTCTCTGCT 3000
ATAAAGACCC TTGCAGCTCT CGTGTTCTGT GAACACTTCC CTGTGATTCT CTGTGAGGGG 3060
GGATGTTGAG AGGGGAAGGA GGCAGAGCTG GAGCAGCTGA GCCACAGGGG AGGTGGAGGG 3120
GGACAGGAAG GCAGGCAGAA GCTGGGTGCT CCATCAGTCC TCACTGATCA CGTCAGACTC 3180
CAGGACCGAG AGCCACAATG CTTCAGGAAA GCTCAATGAA CCCAACAGCC ACATTTTCCT 3240
TCCCTAAGCA TAGACAATGG CATTTGCCAA TAACCAAAAA GAATGCAGAG ACTAACTGGT 3300
GGTAGCTTTT GCCTGGCATT CAAAAACTGG GCCAGAGCAA GTGGAAAATG CCAGAGATTG 3360
TTAAACTTTT CACCCTGACC AGCACCCCAC GCAGCTCAGC AGTGACTGCT GACAGCACGG 3420
AGTGACCTGC AGCGCAGGGG AGGAGAAGAA AAAGAGAGGG ATAGTGTATG AGCAAGAAAG 3480
ACAGATTCAT TCAAGGGCAG TGGGAATTGA CCACAGGGAT TATAGTCCAC GTGATCCTGG 3540
GTTCTAGGAG GCAGGGCTAT ATTGTGGGGG GAAAAAATCA GTTCAAGGGA AGTCGGGAGA 3600
CCTGATTTCT AATACTATAT TTTTCCTTTA CAAGCTGAGT AATTCTGAGC AAGTCACAAG 3660
GTAGTAACTG AGGCTGTAAG ATTACTTAGT TTCTCCTTAT TAGGAACTCT TTTTCTCTGT 3720
GGAGTTAGCA GCACAAGGGC AATCCCGTTT CTTTTAACAG GAAGAAAACA TTCCTAAGAG 3780
TAAAGCCAAA CAGATTCAAG CCTAGGTCTT GCTGACTATA TGATTGGTTT TTTGAAAAAT 3840
CATTTCAGCG ATGTTTACTA TCTGATTCAG AAAATGAGAC TAGTACCCTT TGGTCAGCTG 3900
TAAACAAACA CCCATTTGTA AATGTCTCAA GTTCAGGCTT AACTGCAGAA CCAATCAAAT 3960
AAGAATAGAA TCTTTAGAGC AAACTGTGTT TCTCCACTCT GGAGGTGAGT CTGCCAGGGC 4020

```
AGTTTGGAAA TATTTACTTC ACAAGTATTG ACACTGTTGT TGGTATTAAC AACATAAAGT   4080
TGCTCAAAGG CAATCATTAT TTCAAGTGGC TTAAAGTTAC TTCTGACAGT TTTGGTATAT   4140
TTATTGGCTA TTGCCATTTG CTTTTTGTTT TTTCTCTTTG GGTTTATTAA TGTAAAGCAG   4200
GGATTATTAA CCTACAGTCC AGAAAGCCTG TGAATTTGAA TGAGGAAAAA ATTACATTTT   4260
TGTTTTTACC ACCTTCTAAC TAAATTTAAC ATTTTATTCC ATTGCGAATA GAGCCATAAA   4320
CTCAAAGTGG TAATAACAGT ACCTGTGATT TTGTCATTAC CAATAGAAAT CACAGACATT   4380
TTATACTATA TTACAGTTGT TGCAGATACG TTGTAAGTGA AATATTTATA CTCAAAACTA   4440
CTTTGAAATT AGACCTCCTG CTGGATCTTG TTTTTAACAT ATTAATAAAA CATGTTTAAA   4500
ATTTTGATAT TTTGATAATC ATATTTCATT ATCATTTGTT TCCTTTGTAA TCTATATTTT   4560
ATATATTTGA AAACATCTTT CTGAGAAGAG TTCCCCAGAT TTCACCAATG AGGTTCTTGG   4620
CATGCACACA CACAGAGTAA GAACTGATTT AGAGGCTAAC ATTGACATTG GTGCCTGAGA   4680
TGCAAGACTG AAATTAGAAA GTTCTCCCAA AGATACACAG TTGTTTTAAA GCTAGGGGTG   4740
AGGGGGGAAA TCTGCCGCTT CTATAGGAAT GCTCTCCCTG GAGCCTGGTA GGGTGCTGTC   4800
CTTGTGTTCT GGCTGGCTGT TATTTTTCTC TGTCCCTGCT ACGTCTTAAA GGACTTGTTT   4860
GGATCTCCAG TTCCTAGCAT AGTGCCTGGC ACAGTGCAGG TTCTCAATGA GTTTGCAGAG   4920
TGAATGGAAA TATAAACTAG AAATATATCC TTGTTGAAAT CAGCACACCA GTAGTCCTGG   4980
TGTAAGTGTG TGTACGTGTG TGTGTGTGTG TGTGTGTGTG TGTAAAACCA GGTGGAGATA   5040
TAGGAACTAT TATTGGGGTA TGGGTGCATA AATTGGGATG TTCTTTTTAA AAAGAAACTC   5100
CAAACAGACT TCTGGAAGGT TATTTTCTAA GAATCTTGCT GGCAGCGTGA AGGCAACCCC   5160
CCTGTGCACA GCCCCACCCA GCCTCACGTG GCCACCTCTG TCTTCCCCCA TGAAGGGCTG   5220
GCTCCCCAGT ATATATAAAC CTCTCTGGAG CTCGGGCATG AGCCAGCAAG GCCACCCATC   5280
CAGGCACCTC TCAGCACAGC                                               5300
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5304 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Sub a1 cont

```
ATCTTTGTTC AGTTTACCTC AGGGCTATTA TGAAATGAAA TGAGATAACC AATGTGAAAG     60
TCCTATAAAC TGTATAGCCT CCATTCGGAT GTATGTCTTT GGCAGGATGA TAAAGAATCA    120
GGAAGAAGGA GTATCCACGT TAGCCAAGTG TCCAGGCTGT GTCTGCTCTT ATTTTAGTGA    180
CAGATGTTGC TCCTGACAGA AGCTATTCTT CAGGAAACAT CACATCCAAT ATGGTAAATC    240
CATCAAACAG GAGCTAAGAA ACAGGAATGA GATGGGCACT TGCCCAAGGA AAAATGCCAG    300
GAGAGCAAAT AATGATGAAA AATAAACTTT TCCCTTTGTT TTTAATTTCA GGAAAAAATG    360
ATGAGGACCA AAATCAATGA ATAAGGAAAA CAGCTCAGAA AAAAGATGTT TCCAAATTGG    420
TAATTAAGTA TTTGTTCCTT GGGAAGAGAC CTCCATGTGA GCTTGATGGG AAAATGGGAA    480
AAACGTCAAA AGCATGATCT GATCAGATCC CAAAGTGGAT TATTATTTTA AAAACCAGAT    540
GGCATCACTC TGGGGAGGCA AGTTCAGGAA GGTCATGTTA GCAAAGGACA TAACAATAAC    600
AGCAAAATCA AAATTCCGCA AATGCAGGAG GAAAATGGGG ACTGGGAAAG CTTTCATAAC    660
AGTGATTAGG CAGTTGACCA TGTTCGCAAC ACCTCCCCGT CTATACCAGG GAACACAAAA    720
ATTGACTGGG CTAAGCCTGG ACTTTCAAGG GAAATATGAA AAACTGAGAG CAAAACAAAA    780
GACATGGTTA AAAGGCAACC AGAACATTGT GAGCCTTCAA AGCAGCAGTG CCCCTCAGCA    840
GGGACCCTGA GGCATTTGCC TTTAGGAAGG CCAGTTTTCT TAAGGAATCT TAAGAAACTC    900
TTGAAAGATC ATGAATTTTA ACCATTTTAA GTATAAAACA AATATGCGAT GCATAATCAG    960
TTTAGACATG GGTCCCAATT TTATAAAGTC AGGCATACAA GGATAACGTG TCCCAGCTCC   1020
GGATAGGTCA GAAATCATTA GAAATCACTG TGTCCCCATC CTAACTTTTT CAGAATGATC   1080
TGTCATAGCC CTCACACACA GGCCCGATGT GTCTGACCTA CAACCACATC TACAACCCAA   1140
GTGCCTCAAC CATTGTTAAC GTGTCATCTC AGTAGGTCCC ATTACAAATG CCACCTCCCC   1200
TGTGCAGCCC ATCCCGCTCC ACAGGAAGTC TCCCCACTCT AGACTTCTGC ATCACGATGT   1260
TACAGCCAGA AGCTCCGTGA GGGTGAGGGT CTGTGTCTTA CACCTACCTG TATGCTCTAC   1320
ACCTGAGCTC ACTGCAACCT CTGCCTCCCA GGTTCAAGCA ATTCTCCTGT CTCAGCCTCC   1380
CGCGTAGCTG GGACTACAGG CGCACGCCCG GCTAATTTTT GTATTGTTAG TAGAGATGGG   1440
GTTTCACCAT ATTAGCCCGG CTGGTCTTGA ACTCCTGACC TCAGGTGATC CACCCACCTC   1500
AGCCTCCTAA AGTGCTGGGA TTACAGGCAT GAGTCACCGC GCCCGGCCAA GGGTCAGTGT   1560
TTAATAAGGA ATAACTTGAA TGGTTTACTA AACCAACAGG GAAACAGACA AAAGCTGTGA   1620
TAATTTCAGG GATTCTTGGG ATGGGGAATG GTGCCATGAG CTGCCTGCCT AGTCCCAGAC   1680
CACTGGTCCT CATCACTTTC TTCCCTCATC CTCATTTTCA GGCTAAGTTA CCATTTTATT   1740
CACCATGCTT TTGTGGTAAG CCTCCACATC GTTACTGAAA TAAGAGTATA CATAAACTAG   1800
TTCCATTTGG GGCCATCTGT GTGTGTGTAT AGGGGAGGAG GGCATACCCC AGAGACTCCT   1860
TGAAGCCCCC GGCAGAGGTT TCCTCTCCAG CTGGGGGAGC CCTGCAAGCA CCCGGGGTCC   1920
```

TGGGTGTCCT GAGCAACCTG CCAGCCCGTG CCACTGGTTG TTTTGTTATC ACTCTCTAGG 1980
GACCTGTTGC TTTCTATTTC TGTGTGACTC GTTCATTCAT CCAGGCATTC ATTGACAATT 2040
TATTGAGTAC TTATATCTGC CAGACACCAG AGACAAAATG GTGAGCAAAG CAGTCACTGC 2100
CCTACCTTCG TGGAGGTGAC AGTTTCTCAT GGAAGACGTG CAGAAGAAAA TTAATAGCCA 2160
GCCAACTTAA ACCCAGTGCT GAAAGAAAGG AAATAAACAC CATCTTGAAG AATTGTGCGC 2220
AGCATCCCTT AACAAGGCCA CCTCCCTAGC GCCCCCTGCT GCCTCCATCG TGCCCGGAGG 2280
CCCCCAAGCC CGAGTCTTCC AAGCCTCCTC CTCCATCAGT CACAGCGCTG CAGCTGGCCT 2340
GCCTCGCTTC CCGTGAATCG TCCTGGTGCA TCTGAGCTGG AGACTCCTTG GCTCCAGGCT 2400
CCAGAAAGGA AATGGAGAGG GAAACTAGTC TAACGGAGAA TCTGGAGGGG ACAGTGTTTC 2460
CTCAGAGGGA AAGGGGCCTC CACGTCCAGG AGAATTCCAG GAGGTGGGGA CTGCAGGGAG 2520
TGGGGACGCT GGGGCTGAGC GGGTGCTGAA AGGCAGGAAG GTGAAAAGGG CAAGGCTGAA 2580
GCTGCCCAGA TGTTCAGTGT TGTTCACGGG GCTGGGAGTT TTCCGTTGCT TCCTGTGAGC 2640
CTTTTTATCT TTTCTCTGCT TGGAGGAGAA GAAGTCTATT TCATGAAGGG ATGCAGTTTC 2700
ATAAAGTCAG CTGTTAAAAT TCCAGGGTGT GCATGGGTTT TCCTTCACGA AGGCCTTTAT 2760
TTAATGGGAA TATAGGAAGC GAGCTCATTT CCTAGGCCGT TAATTCACGG AAGAAGTGAC 2820
TGGAGTCTTT TCTTTCATGT CTTCTGGGCA ACTACTCAGC CCTGTGGTGG ACTTGGCTTA 2880
TGCAAGACGG TCGAAAACCT TGGAATCAGG AGACTCGGTT TTCTTTCTGG TTCTGCCATT 2940
GGTTGGCTGT GCGACCGTGG GCAAGTGTCT CTCCTTCCCT GGGCCATAGT CTTCTCTGCT 3000
ATAAAGACCC TTGCAGCTCT CGTGTTCTGT GAACACTTCC CTGTGATTCT CTGTGAGGGG 3060
GGATGTTGAG AGGGGAAGGA GGCAGAGCTG GAGCAGCTGA GCCACAGGGG AGGTGGAGGG 3120
GGACAGGAAG GCAGGCAGAA GCTGGGTGCT CCATCAGTCC TCACTGATCA CGTCAGACTC 3180
CAGGACCGAG AGCCACAATG CTTCAGGAAA GCTCAATGAA CCCAACAGCC ACATTTTCCT 3240
TCCCTAAGCA TAGACAATGG CATTTGCCAA TAACCAAAAA GAATGCAGAG ACTAACTGGT 3300
GGTAGCTTTT GCCTGGCATT CAAAAACTGG GCCAGAGCAA GTGGAAAATG CCAGAGATTG 3360
TTAAACTTTT CACCCTGACC AGCACCCCAC GCAGCTCAGC AGTGACTGCT GACAGCACGG 3420
AGTGACCTGC AGCGCAGGGG AGGAGAAGAA AAAGAGAGGG ATAGTGTATG AGCAAGAAAG 3480
ACAGATTCAT TCAAGGGCAG TGGGAATTGA CCACAGGGAT TATAGTCCAC GTGATCCTGG 3540
GTTCTAGGAG GCAGGGCTAT ATTGTGGGGG GAAAAAATCA GTTCAAGGGA AGTCGGGAGA 3600
CCTGATTTCT AATACTATAT TTTTCCTTTA CAAGCTGAGT AATTCTGAGC AAGTCACAAG 3660
GTAGTAACTG AGGCTGTAAG ATTACTTAGT TTCTCCTTAT TAGGAACTCT TTTTCTCTGT 3720
GGAGTTAGCA GCACAAGGGC AATCCCGTTT CTTTTAACAG GAAGAAAACA TTCCTAAGAG 3780
TAAAGCCAAA CAGATTCAAG CCTAGGTCTT GCTGACTATA TGATTGGTTT TTTGAAAAAT 3840
CATTTCAGCG ATGTTTACTA TCTGATTCAG AAAATGAGAC TAGTACCCTT TGGTCAGCTG 3900
TAAACAAACA CCCATTTGTA AATGTCTCAA GTTCAGGCTT AACTGCAGAA CCAATCAAAT 3960
AAGAATAGAA TCTTTAGAGC AAACTGTGTT TCTCCACTCT GGAGGTGAGT CTGCCAGGGC 4020
AGTTTGGAAA TATTTACTTC ACAAGTATTG ACACTGTTGT TGGTATTAAC AACATAAAGT 4080
TGCTCAAAGG CAATCATTAT TTCAAGTGGC TTAAAGTTAC TTCTGACAGT TTTGGTATAT 4140
TTATTGGCTA TTGCCATTTG CTTTTTGTTT TTTCTCTTTG GGTTTATTAA TGTAAAGCAG 4200
GGATTATTAA CCTACAGTCC AGAAAGCCTG TGAATTTGAA TGAGGAAAAA ATTACGTTTT 4260
TATTTTTACC ACCTTCTAAC TAAATTTAAC ATTTTATTCC ATTGCGAATA GAGCCATAAA 4320
CTCAAAGTGG TAATAAGAGT ACCTGTGATT TTGTCATTAC CAATAGAAAT CACAGACATT 4380
TTATACTATA TTACAGTTGT TGCAGGTACG TTGTAAGTGA AATATTTATA CTCAAAACTA 4440
CTTTGAAATT AGACCTCCTG CTGGATCTTG TTTTTAACAT ATTAATAAAA CATGTTTAAA 4500
ATTTTGATAT TTTGATAATC ATATTTCATT ATCATTTGTT TCCTTTGTAA TCTATATTTT 4560
ATATATTTGA AAACATCTTT CTGAGAAGAG TTCCCCAGAT TTCACCAATG AGGTTCTTGG 4620
CATGCACACA CACAGAGTAA GAACTGATTT AGAGGCTAAC ATTGACATTG GTGCCTGAGA 4680
TGCAAGACTG AAATTAGAAA GTTCTCCCAA AGATACACAG TTGTTTTAAA GCTAGGGGTG 4740
AGGGGGGAAA TCTGCCGCTT CTATAGGAAT GCTCTCCCTG GAGCCTGGTA GGGTGCTGTC 4800
CTTGTGTTCT GGCTGGCTGT TATTTTTCTC TGTCCCTGCT ACGTCTTAAA GGACTTGTTT 4860
GGATCTCCAG TTCCTAGCAT AGTGCCTGGC ACAGTGCAGG TTCTCAATGA GTTTGCAGAG 4920
TGAATGGAAA TATAAACTAG AAATATATCT TTGTTGAAAT CAGCACACCA GTAGTCCTGG 4980
TGTAAGTGTG TGTACGTGTG TGTGTGTGTG TGTGTGTGTG TGTGTGTAAA ACCAGGTGGA 5040
GATATAGGAA CTATTATTGG GGTATGGGTG CATAAATTGG GATGTTCTTT TTAAAAAGAA 5100
ACTCCAAACA GACTTCTGGA AGGTTATTTT CTAAGAATCT TGGTGGCAGC GTGAAGGCAA 5160
CCCCCCTGTG CACAGCCCCA CCCAGCCTCA CGTGGCCACC TCTGTCTTCC CCCATGAAGG 5220
GCTGGCTCCC CAGTATATAT AAACCTCTCT GGAGCTCGGG CATGAGCCAG CAAGGCCACC 5280
CATCCAGGCA CCTCTCAGCA CAGC 5304

Sub a1 cont

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6169 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATCTTTGTTC AGTTTACCTC AGGGCTATTA TGAAATGAAA TGAGATAACC AATGTGAAAG 60
TCCTATAAAC TGTATAGCCT CCATTCGGAT GTATGTCTTT GGCAGGATGA TAAAGAATCA 120
GGAAGAAGGA GTATCCACGT TAGCCAAGTG TCCAGGCTGT GTCTGCTCTT ATTTTAGTGA 180
CAGATGTTGC TCCTGACAGA AGCTATTCTT CAGGAAACAT CACATCCAAT ATGGTAAATC 240
CATCAAACAG GAGCTAAGAA ACAGGAATGA GATGGGCACT TGCCCAAGGA AAAATGCCAG 300
GAGAGCAAAT AATGATGAAA AATAAACTTT TCCCTTTGTT TTTAATTTCA GGAAAAAATG 360
ATGAGGACCA AAATCAATGA ATAAGGAAAA CAGCTCAGAA AAAAGATGTT TCCAAATTGG 420
TAATTAAGTA TTTGTTCCTT GGGAAGAGAC CTCCATGTGA GCTTGATGGG AAAATGGGAA 480
AAACGTCAAA AGCATGATCT GATCAGATCC CAAAGTGGAT TATTATTTTA AAAACCAGAT 540
GGCATCACTC TGGGGAGGCA AGTTCAGGAA GGTCATGTTA GCAAAGGACA TAACAATAAC 600
AGCAAAATCA AAATTCCGCA AATGCAGGAG GAAAATGGGG ACTGGGAAAG CTTTCATAAC 660
AGTGATTAGG CAGTTGACCA TGTTCGCAAC ACCTCCCCGT CTATACCAGG GAACACAAAA 720
ATTGACTGGG CTAAGCCTGG ACTTTCAAGG GAAATATGAA AAACTGAGAG CAAAACAAAA 780
GACATGGTTA AAAGGCAACC AGAACATTGT GAGCCTTCAA AGCAGCAGTG CCCCTCAGCA 840
GGGACCCTGA GGCATTTGCC TTTAGGAAGG CCAGTTTTCT TAAGGAATCT TAAGAAACTC 900
TTGAAAGATC ATGAATTTTA ACCATTTTAA GTATAAAACA AATATGCGAT GCATAATCAG 960
TTTAGACATG GGTCCCAATT TTATAAAGTC AGGCATACAA GGATAACGTG TCCCAGCTCC 1020
GGATAGGTCA GAAATCATTA GAAATCACTG TGTCCCCATC CTAACTTTTT CAGAATGATC 1080
TGTCATAGCC CTCACACACA GGCCCGATGT GTCTGACCTA CAACCACATC TACAACCCAA 1140
GTGCCTCAAC CATTGTTAAC GTGTCATCTC AGTAGGTCCC ATTACAAATG CCACCTCCCC 1200
TGTGCAGCCC ATCCCGCTCC ACAGGAAGTC TCCCCACTCT AGACTTCTGC ATCACGATGT 1260
TACAGCCAGA AGCTCCGTGA GGGTGAGGGT CTGTGTCTTA CACCTACCTG TATGCTCTAC 1320
ACCTGAGCTC ACTGCAACCT CTGCCTCCCA GGTTCAAGCA ATTCTCCTGT CTCAGCCTCC 1380
CGCGTAGCTG GGACTACAGG CGCACGCCCG GCTAATTTTT GTATTGTTAG TAGAGATGGG 1440
GTTTCACCAT ATTAGCCCGG CTGGTCTTGA ACTCCTGACC TCAGGTGATC CACCCACCTC 1500
AGCCTCCTAA AGTGCTGGGA TTACAGGCAT GAGTCACCGC GCCCGGCCAA GGGTCAGTGT 1560
TTAATAAGGA ATAACTTGAA TGGTTTACTA AACCAACAGG GAAACAGACA AAAGCTGTGA 1620
TAATTTCAGG GATTCTTGGG ATGGGGAATG GTGCCATGAG CTGCCTGCCT AGTCCCAGAC 1680
CACTGGTCCT CATCACTTTC TTCCCTCATC CTCATTTTCA GGCTAAGTTA CCATTTTATT 1740
CACCATGCTT TTGTGGTAAG CCTCCACATC GTTACTGAAA TAAGAGTATA CATAAACTAG 1800
TTCCATTTGG GGCCATCTGT GTGTGTGTAT AGGGGAGGAG GGCATACCCC AGAGACTCCT 1860
TGAAGCCCCC GGCAGAGGTT TCCTCTCCAG CTGGGGGAGC CCTGCAAGCA CCCGGGGTCC 1920
TGGGTGTCCT GAGCAACCTG CCAGCCCGTG CCACTGGTTG TTTTGTTATC ACTCTCTAGG 1980
GACCTGTTGC TTTCTATTTC TGTGTGACTC GTTCATTCAT CCAGGCATTC ATTGACAATT 2040
TATTGAGTAC TTATATCTGC CAGACACCAG AGACAAAATG GTGAGCAAAG CAGTCACTGC 2100
CCTACCTTCG TGGAGGTGAC AGTTTCTCAT GGAAGACGTG CAGAAGAAAA TTAATAGCCA 2160
GCCAACTTAA ACCCAGTGCT GAAAGAAAGG AAATAAACAC CATCTTGAAG AATTGTGCGC 2220
AGCATCCCTT AACAAGGCCA CCTCCCTAGC GCCCCCTGCT GCCTCCATCG TGCCCGGAGG 2280
CCCCCAAGCC CGAGTCTTCC AAGCCTCCTC CTCCATCAGT CACAGCGCTG CAGCTGGCCT 2340
GCCTCGCTTC CCGTGAATCG TCCTGGTGCA TCTGAGCTGG AGACTCCTTG GCTCCAGGCT 2400
CCAGAAAGGA AATGGAGAGG GAAACTAGTC TAACGGAGAA TCTGGAGGGG ACAGTGTTTC 2460
CTCAGAGGGA AAGGGGCCTC CACGTCCAGG AGAATTCGAG GAGGTGGGGA CTGCAGGGAG 2520
Sub A1 cont
TGGGGACGCT GGGGCTGAGC GGGTGCTGAA AGGCAGGAAG GTGAAAAGGG CAAGGCTGAA 2580
GCTGCCCAGA TGTTCAGTGT TGTTCACGGG GCTGGGAGTT TTCCGTTGCT TCCTGTGAGC 2640
CTTTTTATCT TTTCTCTGCT TGGAGGAGAA GAAGTCTATT TCATGAAGGG ATGCAGTTTC 2700
ATAAAGTCAG CTGTTAAAAT TCCAGGGTGT GCATGGGTTT TCCTTCACGA AGGCCTTTAT 2760
TTAATGGGAA TATAGGAAGC GAGCTCATTT CCTAGGCCGT TAATTCACGG AAGAAGTGAC 2820
TGGAGTCTTT TCTTTCATGT CTTCTGGGCA ACTACTCAGC CCTGTGGTGG ACTTGGCTTA 2880
TGCAAGACGG TCGAAAACCT TGGAATCAGG AGACTCGGTT TTCTTTCTGG TTCTGCCATT 2940
GGTTGGCTGT GCGACCGTGG GCAAGTGTCT CTCCTTCCCT GGGCCATAGT CTTCTCTGCT 3000
ATAAAGACCC TTGCAGCTCT CGTGTTCTGT GAACACTTCC CTGTGATTCT CTGTGAGGGG 3060
GGATGTTGAG AGGGGAAGGA GGCAGAGCTG GAGCAGCTGA GCCACAGGGG AGGTGGAGGG 3120
GGACAGGAAG GCAGGCAGAA GCTGGGTGCT CCATCAGTCC TCACTGATCA CGTCAGACTC 3180
CAGGACCGAG AGCCACAATG CTTCAGGAAA GCTCAATGAA CCCAACAGCC ACATTTTCCT 3240
TCCCTAAGCA TAGACAATGG CATTTGCCAA TAACCAAAAA GAATGCAGAG ACTAACTGGT 3300
GGTAGCTTTT GCCTGGCATT CAAAAACTGG GCCAGAGCAA GTGGAAAATG CCAGAGATTG 3360
TTAAACTTTT CACCCTGACC AGCACCCCAC GCAGCTCAGC AGTGACTGCT GACAGCACGG 3420
AGTGACCTGC AGCGCAGGGG AGGAGAAGAA AAAGAGAGGG ATAGTGTATG AGCAAGAAAG 3480
ACAGATTCAT TCAAGGGCAG TGGGAATTGA CCACAGGGAT TATAGTCCAC GTGATCCTGG 3540
GTTCTAGGAG GCAGGGCTAT ATTGTGGGGG GAAAAAATCA GTTCAAGGGA AGTCGGGAGA 3600
CCTGATTTCT AATACTATAT TTTTCCTTTA CAAGCTGAGT AATTCTGAGC AAGTCACAAG 3660
GTAGTAACTG AGGCTGTAAG ATTACTTAGT TTCTCCTTAT TAGGAACTCT TTTTCTCTGT 3720
GGAGTTAGCA GCACAAGGGC AATCCCGTTT CTTTTAACAG GAAGAAAACA TTCCTAAGAG 3780

TAAAGCCAAA CAGATTCAAG CCTAGGTCTT GCTGACTATA TGATTGGTTT TTTGAAAAAT 3840
CATTTCAGCG ATGTTTACTA TCTGATTCAG AAAATGAGAC TAGTACCCTT TGGTCAGCTG 3900
TAAACAAACA CCCAGTTGTA AATGTCTCAA GTTCAGGCTT AACTGCAGAA CCAATCAAAA 3960
AGAATAGAAT CTTTAGAGCA AACTGTGTTT CTCCACATCT GGAGGTGAGT CTGCCAGGGC 4020
AGTTTGGAAA TATTTACTTC ACAAGTATTG ACACTGTTGT TGGTATTAAC AACATAAAGT 4080
TGCTCAAAGG CAATCATTAT TTCAAGTGGC TTAAAGTTAC TTCTGACAGT TTTGGTATAT 4140
TTATTGGCTA TTGCCATTTG CTTTTTGTTT TTTCTCTTTG GGTTTATTAA TGTAAAGCAG 4200
GGATTATTAA CCTACAGTCC AGAAAGCCTG TGAATTTGAA TGAGGAAAAA ATTACATTTT 4260
TGTTTTTACC ACCTTCTAAC TAAATTTAAC ATTTTATTCC ATTGCGAATA GAGCCATAAA 4320
CTCAAAGTGG TAATAACAGT ACCTGTGATT TTGTCATTAC CAATAGAAAT CACAGACATT 4380
TTATACTATA TTACAGTTGT TGCAGATACG TTGTAAGTGA AATATTTATA CTCAAAACTA 4440
CTTTGAAATT AGACCTCCTG CTGGATCTTG TTTTTAACAT ATTAATAAAA CATGTTTAAA 4500
ATTTTGATAT TTTGATAATC ATATTTCATT ATCATTTGTT TCCTTTGTAA TCTATATTTT 4560
ATATATTTGA AAACATCTTT CTGAGAAGAG TTCCCCAGAT TTCACCAATG AGGTTCTTGG 4620
CATGCACACA CACAGAGTAA GAACTGATTT AGAGGCTAAC ATTGACATTG GTGCCTGAGA 4680
TGCAAGACTG AAATTAGAAA GTTCTCCCAA AGATACACAG TTGTTTTAAA GCTAGGGGTG 4740
AGGGGGGAAA TCTGCCGCTT CTATAGGAAT GCTCTCCCTG GAGCCTGGTA GGGTGCTGTC 4800
CTTGTGTTCT GGCTGGCTGT TATTTTTCTC TGTCCCTGCT ACGTCTTAAA GGACTTGTTT 4860
GGATCTCCAG TTCCTAGCAT AGTGCCTGGC ACAGTGCAGG TTCTCAATGA GTTTGCAGAG 4920
TGAATGGAAA TATAAACTAG AAATATATCC TTGTTGAAAT CAGCACACCA GTAGTCCTGG 4980
TGTAAGTGTG TGTACGTGTG TGTGTGTGTG TGTGTGTGTG TGTAAAACCA GGTGGAGATA 5040
TAGGAACTAT TATTGGGGTA TGGGTGCATA AATTGGGATG TTCTTTTTAA AAAGAAACTC 5100
CAAACAGACT TCTGGAAGGT TATTTTCTAA GAATCTTGCT GGCAGCGTGA AGGCAACCCC 5160
CCTGTGCACA GCCCCACCCA GCCTCACGTG GCCACCTCTG TCTTCCCCCA TGAAGGGCTG 5220
GCTCCCCAGT ATATATAAAC CTCTCTGGAG CTCGGGCATG AGCCAGCAAG GCCACCCATC 5280
CAGGCACCTC TCAGCACAGC AGAGCTTTCC AGAGGAAGCC TCACCAAGCC TCTGCAATGA 5340
GGTTCTTCTG TGCACGTTGC TGCAGCTTTG GGCCTGAGAT GCCAGCTGTC CAGCTGCTGC 5400
TTCTGGCCTG CCTGGTGTGG GATGTGGGGG CCAGGACAGC TCAGCTCAGG AAGGCCAATG 5460
ACCAGAGTGG CCGATGCCAG TATACCTTCA GTGTGGCCAG TCCCAATGAA TCCAGCTGCC 5520
CAGAGCAGAG CCAGGCCATG TCAGTCATCC ATAACTTACA GAGAGACAGC AGCACCCAAC 5580
GCTTAGACCT GGAGGCCACC AAAGCTCGAC TCAGCTCCCT GGAGAGCCTC CTCCACCAAT 5640
TGACCTTGGA CCAGGCTGCC AGGCCCCAGG AGACCCAGGA GGGGCTGCAG AGGGAGCTGG 5700
GCACCCTGAG GCGGGAGCGG GACCAGCTGG AAACCCAAAC CAGAGAGTTG GAGACTGCCT 5760
ACAGCAACCT CCTCCGAGAC AAGTCAGTTC TGGAGGAAGA GAAGAAGCGA CTAAGGCAAG 5820
AAAATGAGAA TCTGGCCAGG AGGTTGGAAA GCAGCAGCCA GGAGGTAGCA AGGCTGAGAA 5880
GGGGCCAGTG TCCCCAGACC CGAGACACTG CTCGGGCTGT GCCACCAGGC TCCAGAGAAG 5940
GTAAGAATGC AGAGTGGGGG GACTCTGAGT TCAGCAGGTG ATATGGCTCG TAGTGACCTG 6000
CTACAGGCGC TCCAGGCCTC CCTGCCCTTT CTCCTAGAGA CTGCACAGCT AGCACAAGAC 6060
AGATGAATTA AGGAAAGCAC ACGATCACCT TCAAGTATTA CTAGTAATTT AGCTCCTGAG 6120
AGCTTCATTT AGATTAGTGG TTCAGAGTTC TTGTGCCCCT CCATGTCAG 6169

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 926 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

Sub a1 cont

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AAGGTAGGCA CATTGCCCTG CAATTTATAA TTTATGAGGT GTTCAATTAT GGAATTGTCA 60
AATATTAACA AAAGTAGAGA GACTACAATG AACTCCAATG TAGCCATAAC TCAGGCCCAA 120
CTGTTATCAG CACAGTCCAA TCATGTTTTA TCTTTCCTTC TCTGACCCCC AACCCATCCC 180
CAGTCCTTAT CTAAAATCAA ATATCAAACA CCATACTCTT TGGGAGCCTA TTTATTTAGT 240
TAGTTAGTTT TCAGACAGAG TTTCTTTCTT GTTCCCAAGC TGGAGTACAA TAGTGTAGTC 300
TCGGCTAACA GCAATCTCCC CCTCCTTGGT TCAAGCAATT CTCCTGCCTC AGTCTCCCAA 360
GAAGCTGGGA TTATAGACAC CTGCCACCAC ATCCAGCTAA TTTTTTTGTG TTTTAGAAAA 420
GACAGGGTTT CACCATGTTG GCCAGGCTGG TTTCGAACTC CTGACCTCAG GTGATCCGCC 480
TGCCTCGGCC TCCCAAAGTG CTGGGATTAC AGGCATGAGC CACCACGCCT GGCCGGCAGC 540
CTATTTAAAT GTCATCCTCA ACATAGTCAA TCCTTGGGCC ATTTTTTCTT ACAGTAAAAT 600
TTTGTCTCTT TCTTTTAATC AGTTTCTACG TGGAATTTGG ACACTTTGGC CTTCCAGGAA 660
CTGAAGTCCG AGCTAACTGA AGTTCCTGCT TCCCGAATTT TGAAGGAGAG CCCATCTGGC 720
TATCTCAGGA GTGGAGAGGG AGACACCGGT ATGAAGTTAA GTTTCTTCCC TTTTGTGCCC 780
ACGTGGTCTT TATTCATGTC TAGTGCTGTG TTCAGAGAAT CAGTATAGGG TAAATGCCCA 840

```
CCCAAGGGGG AAATTAACTT CCCTGGGAGC AGAGGGAGGG GAGGAGAAGA GGAACAGAAC     900
TCTCTCTCTC TCTCTGTTAC CCTTGT                                          926
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2099 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
TGGCTCTGCC AAGCTTCCGC ATGATCATTG TCTGTGTTTG GAAGATTATG GATTAAGTGG      60
TGCTTCGTTT TCTTTCTGAA TTTACCAGGA TGTGGAGAAC TAGTTTGGGT AGGAGAGCCT     120
CTCACGCTGA GAACAGCAGA AACAATTACT GGCAAGTATG GTGTGTGGAT GCGAGACCCC     180
AAGCCCACCT ACCCCTACAC CCAGGAGACC ACGTGGAGAA TCGACACAGT TGGCACGGAT     240
GTCCGCCAGG TTTTTGAGTA TGACCTCATC AGCCAGTTTA TGCAGGGCTA CCCTTCTAAG     300
GTTCACATAC TGCCTAGGCC ACTGGAAAGC ACGGGTGCTG TGGTGTACTC GGGGAGCCTC     360
TATTTCCAGG GCGCTGAGTC CAGAACTGTC ATAAGATATG AGCTGAATAC CGAGACAGTG     420
AAGGCTGAGA AGGAAATCCC TGGAGCTGGC TACCACGGAC AGTTCCCGTA TTCTTGGGGT     480
GGCTACACGG ACATTGACTT GGCTGTGGAT GAAGCAGGCC TCTGGGTCAT TTACAGCACC     540
GATGAGGCCA AAGGTGCCAT TGTCCTCTCC AAACTGAACC CAGAGAATCT GGAACTCGAA     600
CAAACCTGGG AGACAAACAT CCGTAAGCAG TCAGTCGCCA ATGCCTTCAT CATCTGTGGC     660
ACCTTGTACA CCGTCAGCAG CTACACCTCA GCAGATGCTA CCGTCAACTT TGCTTATGAC     720
ACAGGCACAG GTATCAGCAA GACCCTGACC ATCCCATTCA AGAACCGCTA TAAGTACAGC     780
AGCATGATTG ACTACAACCC CCTGGAGAAG AAGCTCTTTG CCTGGGACAA CTTGAACATG     840
GTCACTTATG ACATCAAGCT CTCCAAGATG TGAAAAGCCT CCAAGCTGTA CAGGCAATGG     900
CAGAAGGAGA TGCTCAGGGC TCCTGGGGGG AGCAGGCTGA AGGGAGAGCC AGCCAGCCAG     960
GGCCCAGGCA GCTTTGACTG CTTTCCAAGT TTTCATTAAT CCAGAAGGAT GAACATGGTC    1020
ACCATCTAAC TATTCAGGAA TTGTAGTCTG AGGGCGTAGA CAATTTCATA TAATAAATAT    1080
CCTTTATCTT CTGTCAGCAT TTATGGGATG TTTAATGACA TAGTTCAAGT TTTCTTGTGA    1140
TTTGGGGCAA AAGCTGTAAG GCATAATAGT CTTTTCCTGA AAACCATTGC TCTTGCATGT    1200
TACATGGTTA CCACAAGCCA CAATAAAAAG CATAACTTCT AAAGGAAGCA GAATAGCTCC    1260
TCTGGCCAGC ATCGAATATA AGTAAGATGC ATTTACTACA GTTGGCTTCT AATGCTTCAG    1320
ATAGAATACA GTTGGGTCTC ACATAACCCT TACATTGTGA AATAAAATTT TCTTACCCAA    1380
CGTTCTCTTC CTTGAACTTT GTGGGAATCT TTGCTTAAGA GAAGGATATA GATTCCAACC    1440
ATCAGGTAAT TCCTTCAGGT TGGGAGATGT GATTGCAGGA TGTTAAAGGT GTGTGTGTGT    1500
GTGTGTGTGT GTGTGTAACT GAGAGGCTTG TGCCTGGTTT TGAGGTGCTG CCCAGGATGA    1560
CGCCAAGCAA ATAGCGCATC CACACTTTCC CACCTCCATC TCCTGGTGCT CTCGGCACTA    1620
CCGGAGCAAT CTTTCCATCT CTCCCCTGAA CCCACCCTCT ATTCACCCTA ACTCCACTTC    1680
AGTTTGCTTT TGATTTTTTT TTTTTTTTTT TTTTTTTTTT GAGATGGGGT CTCGCTCTGT    1740
CACCCAGGCT GGAGTGCAGT GGCACGATCT CGGCTCACTG CAAGTTCCGC CTCCCAGGTT    1800
CACACCATTC TCCTGCCTCA GCCTCCCAAG TAGCTGGGAC TACAGGCACC TGCCACCACG    1860
CCTGGCTAAT TTTTTTTTTT TCCAGTGAAG ATGGGTTTCA CCATGTTAGC CAGGATGGTC    1920
TCGATCTCCT GACCTTGTCA TCCACCCACC TTGGCCTCCC AAAGTGCTGG GATTACAGGC    1980
GTGAGCCACC ACGCCCAGCC CCTCCACTTC AGTTTTTATC TGTCATCAGG GGTATGAATT    2040
TTATAAGCCA CACCTCAGGT GGAGAAAGCT TGATGCATAG CTTGAGTATT CTATACTGT     2099
```

Sub a1 cont

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
TGAGGCTTCC TCTGGAAAC                                                   19
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 20 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TGAAATCAGC ACACCAGTAG 20

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 21 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GCACCCATAC CCCAATAATA G 21

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 20 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AGAGTTCCCC AGATTTCACC 20

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 20 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

Sub a1 cont

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ATCTGGGGAA CTCTTCTCAG 20

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 19 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TACAGTTGTT GCAGATACG 19

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

ACAACGTATC TGCAACAACT G 21

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TCAGGCTTAA CTGCAGAACC 20

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TTGGTTCTGC AGTTAAGCC 19

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Sub A1 cont

AGCAGCACAA GGGCAATCC 19

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ACAGGGCTAT ATTGTGGG 18

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CCTGAGATGC CAGCTGTCC 19

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CTGAAGCATT AGAAGCCAAC 20

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

ACCTTGGACC AGGCTGCCAG 20

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

AGGTTTGTTC GAGTTCCAG 19

Sub a1 cont.

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ACAATTACTG GCAAGTATGG 20

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CCTTCTCAGC CTTGCTACC 19

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

ACACCTCAGC AGATGCTACC 20

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

ATGGATGACT GACATGGCC 19

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

AAGGATGAAC ATGGTCACC 19

Sub a1 cont

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1548 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AGAGCTTTCC AGAGGAAGCC TCACCAAGCC TCTGCAATGA GGTTCTTCTG TGCACGTTGC 60
TGCAGCTTTG GGCCTGAGAT GCCAGCTGTC CAGCTGCTGC TTCTGGCCTG CCTGGTGTGG 120
GATGTGGGGG CCAGGACAGC TCAGCTCAGG AAGGCCAATG ACCAGAGTGG CCGATGCCAG 180
TATACCTTCA GTGTGGCCAG TCCCAATGAA TCCAGCTGCC CAGAGCAGAG CCAGGCCATG 240
TCAGTCATCC ATAACTTACA GAGAGACAGC AGCACCCAAC GCTTAGACCT GGAGGCCACC 300
AAAGCTCGAC TCAGCTCCCT GGAGAGCCTC CTCCACCAAT TGACCTTGGA CCAGGCTGCC 360
AGGCCCCAGG AGACCCAGGA GGGGCTGCAG AGGGAGCTGG GCACCCTGAG GCGGGAGCGG 420

GACCAGCTGG AAACCCAAAC CAGAGAGTTG GAGACTGCCT ACAGCAACCT CCTCCGAGAC 480
AAGTCAGTTC TGGAGGAAGA GAAGAAGCGA CTAAGGCAAG AAAATGAGAA TCTGGCCAGG 540
AGGTTGGAAA GCAGCAGCCA GGAGGTAGCA AGGCTGAGAA GGGGCCAGTG TCCCCAGACC 600
CGAGACACTG CTCGGGCTGT GCCACCAGGC TCCAGAGAAG TTTCTACGTG GAATTTGGAC 660
ACTTTGGCCT TCCAGGAACT GAAGTCCGAG CTAACTGAAG TTCCTGCTTC CCGAATTTTG 720
AAGGAGAGCC CATCTGGCTA TCTCAGGAGT GGAGAGGGAG ACACCGGATG TGGAGAACTA 780
GTTTGGGTAG GAGAGCCTCT CACGCTGAGA ACAGCAGAAA CAATTACTGG CAAGTATGGT 840
GTGTGGATGC GAGACCCCAA GCCCACCTAC CCCTACACCC AGGAGACCAC GTGGAGAATC 900
GACACAGTTG GCACGGATGT CCGCCAGGTT TTTGAGTATG ACCTCATCAG CCAGTTTATG 960
CAGGGCTACC CTTCTAAGGT TCACATACTG CCTAGGCCAC TGGAAAGCAC GGGTGCTGTG 1020
GTGTACTCGG GGAGCCTCTA TTTCCAGGGC GCTGAGTCCA GAACTGTCAT AAGATATGAG 1080
CTGAATACCG AGACAGTGAA GGCTGAGAAG GAAATCCCTG GAGCTGGCTA CCACGGACAG 1140
TTCCCGTATT CTTGGGGTGG CTACACGGAC ATTGACTTGG CTGTGGATGA AGCAGGCCTC 1200
TGGGTCATTT ACAGCACCGA TGAGGCCAAA GGTGCCATTG TCCTCTCCAA ACTGAACCCA 1260
GAGAATCTGG AACTCGAACA AACCTGGGAG ACAAACATCC GTAAGCAGTC AGTCGCCAAT 1320
GCCTTCATCA TCTGTGGCAC CTTGTACACC GTCAGCAGCT ACACCTCAGC AGATGCTACC 1380
GTCAACTTTG CTTATGACAC AGGCACAGGT ATCAGCAAGA CCCTGACCAT CCCATTCAAG 1440
AACCGCTATA AGTACAGCAG CATGATTGAC TACAACCCCC TGGAGAAGAA GCTCTTTGCC 1500
TGGGACAACT TGAACATGGT CACTTATGAC ATCAAGCTCT CCAAGATG 1548

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 178 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Thr Gly Ala Val Val Tyr Ser Gly Ser Leu Tyr Phe Gln Gly Ala Glu
1 5 10 15
Ser Arg Thr Val Ile Arg Tyr Glu Leu Asn Thr Glu Thr Val Lys Ala
20 25 30
Glu Lys Glu Ile Pro Gly Ala Gly Tyr His Gly Gln Phe Pro Tyr Ser
35 40 45
Trp Gly Gly Tyr Thr Asp Ile Asp Leu Ala Val Asp Glu Ala Gly Leu
50 55 60
Trp Val Ile Tyr Ser Thr Asp Glu Ala Lys Gly Ala Ile Val Leu Ser
65 70 75 80
Lys Leu Asn Pro Glu Asn Leu Glu Leu Glu Gln Thr Trp Glu Thr Asn
85 90 95
Ile Arg Lys Gln Ser Val Ala Asn Ala Phe Ile Ile Cys Gly Thr Leu
100 105 110
Tyr Thr Val Ser Ser Tyr Thr Ser Ala Asp Ala Thr Val Asn Phe Ala
115 120 125
Tyr Asp Thr Gly Thr Gly Ile Ser Lys Thr Leu Thr Ile Pro Phe Lys
130 135 140
Asn Arg Tyr Lys Tyr Ser Ser Met Ile Asp Tyr Asn Pro Leu Glu Lys
145 150 155 160
Lys Leu Phe Ala Trp Asp Asn Leu Asn Met Val Thr Tyr Asp Ile Lys
165 170 175
Leu Ser

Sub a' cont

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 131 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

```
Arg Phe Asp Leu Lys Thr Glu Thr Ile Leu Lys Thr Arg Ser Leu Asp
 1               5                  10                  15
Tyr Ala Gly Tyr Asn Asn Met Tyr His Tyr Ala Trp Gly Gly His Ser
            20                  25                  30
Asp Ile Asp Leu Met Val Asp Glu Ser Gly Leu Trp Ala Val Tyr Ala
        35                  40                  45
Thr Asn Gln Asn Ala Gly Asn Ile Val Val Ser Arg Leu Asp Pro Val
    50                  55                  60
Ser Leu Gln Thr Leu Gln Thr Trp Asn Thr Ser Tyr Pro Lys Arg Xaa
65                  70                  75                  80
Pro Gly Xaa Ala Phe Ile Ile Cys Gly Thr Cys Tyr Val Thr Asn Gly
                85                  90                  95
Tyr Ser Gly Gly Thr Lys Val His Tyr Ala Tyr Gln Thr Asn Ala Ser
            100                 105                 110
Thr Tyr Glu Tyr Ile Asp Ile Pro Phe Gln Asn Lys Leu Xaa Pro His
        115                 120                 125
Phe Pro Cys
    130
```

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 178 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```
Gly Thr Gly Gln Val Val Tyr Asn Gly Ser Ile Tyr Phe Asn Lys Phe
 1               5                  10                  15
Gln Ser His Ile Ile Ile Arg Phe Asp Leu Lys Thr Glu Thr Ile Leu
            20                  25                  30
Lys Thr Arg Ser Leu Asp Tyr Ala Gly Tyr Asn Asn Met Tyr His Tyr
        35                  40                  45
Ala Trp Gly Gly His Ser Asp Ile Asp Leu Met Val Asp Glu Asn Gly
    50                  55                  60
Leu Trp Ala Val Tyr Ala Thr Asn Gln Asn Ala Gly Asn Ile Val Ile
65                  70                  75                  80
Ser Lys Leu Asp Pro Val Ser Leu Gln Ile Leu Gln Thr Trp Asn Thr
                85                  90                  95
Ser Tyr Pro Lys Arg Ser Ala Gly Glu Ala Phe Ile Ile Cys Gly Thr
            100                 105                 110
Leu Tyr Val Thr Asn Gly Tyr Ser Gly Gly Thr Lys Val His Tyr Ala
        115                 120                 125
Tyr Gln Thr Asn Ala Ser Thr Tyr Glu Tyr Ile Asp Ile Pro Phe Gln
    130                 135                 140
Asn Lys Tyr Ser His Ile Ser Met Leu Asp Tyr Asn Pro Lys Asp Arg
145                 150                 155                 160
Ala Leu Tyr Ala Trp Asn Asn Gly His Gln Thr Leu Tyr Asn Val Thr
                165                 170                 175
Leu Phe
```

Sub A1 cont

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 177 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Gly Ala Gly Val Val Val His Asn Asn Asn Leu Tyr Tyr Asn Cys Ph
1 5 10 15
Asn Ser His Asp Met Cys Arg Ala Ser Leu Thr Ser Gly Val Tyr Gln
20 25 30
Lys Lys Pro Leu Leu Asn Ala Leu Phe Asn Asn Arg Phe Ser Tyr Ala
35 40 45
Gly Thr Met Phe Gln Asp Met Asp Phe Ser Ser Asp Glu Lys Gly Leu
50 55 60
Trp Val Ile Phe Thr Thr Glu Lys Ser Ala Gly Lys Ile Val Val Gly
65 70 75 80
Lys Val Asn Val Ala Thr Phe Thr Val Asp Asn Ile Trp Ile Thr Thr
85 90 95
Gln Asn Lys Ser Asp Ala Ser Asn Ala Phe Met Ile Cys Gly Val Leu
100 105 110
Tyr Val Thr Arg Ser Leu Gly Pro Lys Met Glu Glu Val Phe Tyr Met
115 120 125
Phe Asp Thr Lys Thr Gly Lys Glu Gly His Leu Ser Ile Met Met Glu
130 135 140
Lys Met Ala Glu Lys Val His Ser Leu Ser Tyr Asn Ser Asn Asp Arg
145 150 155 160
Lys Leu Tyr Met Phe Ser Glu Gly Tyr Leu Leu His Tyr Asp Ile Ala
165 170 175
Leu

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 74 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Gly Val Val Tyr Ser Arg Leu Thr Glu Thr Leu Ala Gly Tyr Asn Asn
1 5 10 15
Tyr Ala Trp Gly Gly Asp Ile Asp Leu Val Asp Glu Gly Leu Trp Tyr
20 25 30
Thr Ala Gly Ile Val Ser Lys Leu Pro Leu Gln Thr Trp Thr Lys Ala
35 40 45
Phe Ile Ile Cys Gly Thr Leu Tyr Val Thr Tyr Val Tyr Ala Tyr Thr
50 55 60
Ile Tyr Asp Tyr Asn Pro Lys Leu Tyr Leu
65 70

Sub A1 cont

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 504 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: N-t rminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

```
Met Arg Phe Ph  Cys Ala Arg Cys Cys Ser Phe Gly Pro Glu Met Pro
 1               5                   10                  15
Ala Val Gln Leu Leu Leu Leu Ala Cys Leu Val Trp Asp Val Gly Ala
             20                  25                  30
Arg Thr Ala Gln Leu Arg Lys Ala Asn Asp Gln Ser Gly Arg Cys Gln
35                 40                   45
Tyr Thr Phe Ser Val Ala Ser Pro Asn Glu Ser Ser Cys Pro Glu Gln
     50                  55                  60
Ser Gln Ala Met Ser Val Ile His Asn Leu Gln Arg Asp Ser Ser Thr
 65                  70                  75                  80
Gln Arg Leu Asp Leu Glu Ala Thr Lys Ala Arg Leu Ser Ser Leu Glu
                 85                  90                  95
Ser Leu Leu His Gln Leu Thr Leu Asp Gln Ala Ala Arg Pro Gln Glu
             100                 105                 110
Thr Gln Glu Gly Leu Gln Arg Glu Leu Gly Thr Leu Arg Arg Glu Arg
         115                 120                 125
Asp Gln Leu Glu Thr Gln Thr Arg Glu Leu Glu Thr Ala Tyr Ser Asn
     130                 135                 140
Leu Leu Arg Asp Lys Ser Val Leu Glu Glu Glu Lys Lys Arg Leu Arg
 145                 150                 155                 160
Gln Glu Asn Glu Asn Leu Ala Arg Arg Leu Glu Ser Ser Ser Gln Glu
                 165                 170                 175
Val Ala Arg Leu Arg Arg Gly Gln Cys Pro Gln Thr Arg Asp Thr Ala
             180                 185                 190
Arg Ala Val Pro Pro Gly Ser Arg Glu Val Ser Thr Trp Asn Leu Asp
         195                 200                 205
Thr Leu Ala Phe Gln Glu Leu Lys Ser Glu Leu Thr Glu Val Pro Ala
     210                 215                 220
Ser Arg Ile Leu Lys Glu Ser Pro Ser Gly Tyr Leu Arg Ser Gly Glu
 225                 230                 235                 240
Gly Asp Thr Gly Cys Gly Glu Leu Val Trp Val Gly Glu Pro Leu Thr
                 245                 250                 255
Leu Arg Thr Ala Glu Thr Ile Thr Gly Lys Tyr Gly Val Trp Met Arg
             260                 265                 270
Asp Pro Lys Pro Thr Tyr Pro Tyr Thr Gln Glu Thr Thr Trp Arg Ile
         275                 280                 285
Asp Thr Val Gly Thr Asp Val Arg Gln Val Phe Glu Tyr Asp Leu Ile
     290                 295                 300
Ser Gln Phe Met Gln Gly Tyr Pro Ser Lys Val His Ile Leu Pro Arg
 305                 310                 315                 320
Pro Leu Glu Ser Thr Gly Ala Val Val Tyr Ser Gly Ser Leu Tyr Phe
                 325                 330                 335
Gln Gly Ala Glu Ser Arg Thr Val Ile Arg Tyr Glu Leu Asn Thr Glu
             340                 345                 350
Thr Val Lys Ala Glu Lys Glu Ile Pro Gly Ala Gly Tyr His Gly Gln
         355                 360                 365
Phe Pro Tyr Ser Trp Gly Gly Tyr Thr Asp Ile Asp Leu Ala Val Asp
     370                 375                 380
Glu Ala Gly Leu Trp Val Ile Tyr Ser Thr Asp Glu Ala Lys Gly Ala
 385                 390                 395                 400
Ile Val Leu Ser Lys Leu Asn Pro Glu Asn Leu Glu Leu Glu Gln Thr
                 405                 410                 415
Trp Glu Thr Asn Ile Arg Lys Gln Ser Val Ala Asn Ala Phe Ile Ile
             420                 425                 430
Cys Gly Thr Leu Tyr Thr Val Ser Ser Tyr Thr Ser Ala Asp Ala Thr
         435                 440                 445
Val Asn Phe Ala Tyr Asp Thr Gly Thr Gly Ile Ser Lys Thr Leu Thr
     450                 455                 460
Ile Pro Phe Lys Asn Arg Tyr Lys Tyr Ser Ser Met Ile Asp Tyr Asn
 465                 470                 475                 480
Pro Leu Glu Lys Lys Leu Phe Ala Trp Asp Asn Leu Asn Met Val Thr
                 485                 490                 495
Tyr Asp Ile Lys Leu Ser Lys M t
             500
```

Sub a1 cont

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CAAACAGACT TCCGGAAGGT

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5271 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Sub a1 cont.

```
ATCTTTGTTC AGTTTACCTC AGGGCTATTA TGAAATGAAA TGAGATAACC AATGTGAAAG    60
TCCTATAAAC TGTATAGCCT CCATTCGGAT GTATGTCTTT GGCAGGATGA TAAAGAATCA   120
GGAAGAAGGA GTATCCACGT TAGCCAAGTG TCCAGGCTGT GTCTGCTCTT ATTTTAGTGA   180
CAGATGTTGC TCCTGACAGA AGCTATTCTT CAGGAAACAT CACATCCAAT ATGGTAAATC   240
CATCAAACAG GAGCTAAGAA ACAGGAATGA GATGGGCACT TGCCCAAGGA AAAATGCCAG   300
GAGAGCAAAT AATGATGAAA AATAAACTTT TCCCTTTGTT TTTAATTTCA GGAAAAAATG   360
ATGAGGACCA AAATCAATGA ATAAGGAAAA CAGCTCAGAA AAAAGATGTT TCCAAATTGG   420
TAATTAAGTA TTTGTTCCTT GGGAAGAGAC CTCCATGTGA GCTTGATGGG AAAATGGGAA   480
AAACGTCAAA AGCATGATCT GATCAGATCC CAAAGTGGAT TATTATTTTA AAAACCAGAT   540
GGCATCACTC TGGGGAGGCA AGTTCAGGAA GGTCATGTTA GCAAAGGACA TAACAATAAC   600
AGCAAAATCA AAATTCCGCA AATGCAGGAG GAAAATGGGG ACTGGGAAAG CTTTCATAAC   660
AGTGATTAGG CAGTTGACCA TGTTCGCAAC ACCTCCCCGT CTATACCAGG GAACACAAAA   720
ATTGACTGGG CTAAGCCTGG ACTTTCAAGG GAAATATGAA AAACTGAGAG CAAAACAAAA   780
GACATGGTTA AAAGGCAACC AGAACATTGT GAGCCTTCAA AGCAGCAGTG CCCCTCAGCA   840
GGGACCCTGA GGCATTTGCC TTTAGGAAGG CCAGTTTTCT TAAGGAATCT TAAGAAACTC   900
TTGAAAGATC ATGAATTTTA ACCATTTTAA GTATAAAACA AATATGCGAT GCATAATCAG   960
TTTAGACATG GGTCCCAATT TTATAAAGTC AGGCATACAA GGATAACGTG TCCCAGCTCC  1020
GGATAGGTCA GAAATCATTA GAAATCACTG TGTCCCCATC CTAACTTTTT CAGAATGATC  1080
TGTCATAGCC CTCACACACA GGCCCGATGT GTCTGACCTA CAACCACATC TACAACCCAA  1140
GTGCCTCAAC CATTGTTAAC GTGTCATCTC AGTAGGTCCC ATTACAAATG CCACCTCCCC  1200
TGTGCAGCCC ATCCCGCTCC ACAGGAAGTC TCCCCACTCT AGACTTCTGC ATCACGATGT  1260
TACAGCCAGA AGCTCCGTGA GGGTGAGGGT CTGTGTCTTA CACCTACCTG TATGCTCTAC  1320
ACCTGAGCTC ACTGCAACCT CTGCCTCCCA GGTTCAAGCA ATTCTCCTGT CTCAGCCTCC  1380
CGCGTAGCTG GGACTACAGG CGCACGCCCG GCTAATTTTT GTATTGTTAG TAGAGATGGG  1440
GTTTCACCAT ATTAGCCCGG CTGGTCTTGA ACTCCTGACC TCAGGTGATC CACCCACCTC  1500
AGCCTCCTAA AGTGCTGGGA TTACAGGCAT GAGTCACCGC GCCCGGCCAA GGGTCAGTGT  1560
TTAATAAGGA ATAACTTGAA TGGTTTACTA AACCAACAGG GAAACAGACA AAAGCTGTGA  1620
TAATTTCAGG GATTCTTGGG ATGGGGAATG GTGCCATGAG CTGCCTGCCT AGTCCCAGAC  1680
CACTGGTCCT CATCACTTTC TTCCCTCATC CTCATTTTCA GGCTAAGTTA CCATTTTATT  1740
CACCATGCTT TTGTGGTAAG CCTCCACATC GTTACTGAAA TAAGAGTATA CATAAACTAG  1800
TTCCATTTGG GGCCATCTGT GTGTGTGTAT AGGGGAGGAG GGCATACCCC AGAGACTCCT  1860
TGAAGCCCCC GGCAGAGGTT TCCTCTCCAG CTGGGGGAGC CCTGCAAGCA CCCGGGGTCC  1920
TGGGTGTCCT GAGCAACCTG CCAGCCCGTG CCACTGGTTG TTTTGTTATC ACTCTCTAGG  1980
GACCTGTTGC TTTCTATTTC TGTGTGACTC GTTCATTCAT CCAGGCATTC ATTGACAATT  2040
TATTGAGTAC TTATATCTGC CAGACACCAG AGACAAAATG GTGAGCAAAG CAGTCACTGC  2100
```

```
CCTACCTTCG TGGAGGTGAC AGTTTCTCAT GGAAGACGTG CAGAAGAAAA TTAATAGCCA   2160
GCCAACTTAA ACCCAGTGCT GAAAGAAAGG AAATAAACAC CATCTTGAAG AATTGTGCGC   2220
AGCATCCCTT AACAAGGCCA CCTCCCTAGC GCCCCCTGCT GCCTCCATCG TGCCCGGAGG   2280
CCCCCAAGCC CGAGTCTTCC AAGCCTCCTC CTCCATCAGT CACAGCGCTG CAGCTGGCCT   2340
GCCTCGCTTC CCGTGAATCG TCCTGGTGCA TCTGAGCTGG AGACTCCTTG GCTCCAGGCT   2400
CCAGAAAGGA AATGGAGAGG GAAACTAGTC TAACGGAGAA TCTGGAGGGG ACAGTGTTTC   2460
CTCAGAGGGA AAGGGGCCTC CACGTCCAGG AGAATTCCAG GAGGTGGGGA CTGCAGGGAG   2520
TGGGGACGCT GGGGCTGAGC GGGTGCTGAA AGGCAGGAAG GTGAAAAGGG CAAGGCTGAA   2580
GCTGCCCAGA TGTTCAGTGT TGTTCACGGG GCTGGGAGTT TTCCGTTGCT TCCTGTGAGC   2640
CTTTTTATCT TTTCTCTGCT TGGAGGAGAA GAAGTCTATT TCATGAAGGG ATGCAGTTTC   2700
ATAAAGTCAG CTGTTAAAAT TCCAGGGTGT GCATGGGTTT TCCTTCACGA AGGCCTTTAT   2760
TTAATGGGAA TATAGGAAGC GAGCTCATTT CCTAGGCCGT TAATTCACGG AAGAAGTGAC   2820
TGGAGTCTTT TCTTTCATGT CTTCTGGGCA ACTACTCAGC CCTGTGGTGG ACTTGGCTTA   2880
TGCAAGACGG TCGAAAACCT TGGAATCAGG AGACTCGGTT TTCTTTCTGG TTCTGCCATT   2940
GGTTGGCTGT GCGACCGTGG GCAAGTGTCT CTCCTTCCCT GGGCCATAGT CTTCTCTGCT   3000
ATAAAGACCC TTGCAGCTCT CGTGTTCTGT GAACACTTCC CTGTGATTCT CTGTGAGGGG   3060
GGATGTTGAG AGGGGAAGGA GGCAGAGCTG GAGCAGCTGA GCCACAGGGG AGGTGGAGGG   3120
GGACAGGAAG GCAGGCAGAA GCTGGGTGCT CCATCAGTCC TCACTGATCA CGTCAGACTC   3180
CAGGACCGAG AGCCACAATG CTTCAGGAAA GCTCAATGAA CCCAACAGCC ACATTTTCCT   3240
TCCCTAAGCA TAGACAATGG CATTTGCCAA TAACCAAAAA GAATGCAGAG ACTAACTGGT   3300
GGTAGCTTTT GCCTGGCATT CAAAAACTGG GCCAGAGCAA GTGGAAAATG CCAGAGATTG   3360
TTAAACTTTT CACCCTGACC AGCACCCCAC GCAGCTCAGC AGTGACTGCT GACAGCACGG   3420
AGTGACCTGC AGCGCAGGGG AGGAGAAGAA AAAGAGAGGG ATAGTGTATG AGCAAGAAAG   3480
ACAGATTCAT TCAAGGGCAG TGGGAATTGA CCACAGGGAT TATAGTCCAC GTGATCCTGG   3540
GTTCTAGGAG GCAGGGCTAT ATTGTGGGGG GAAAAAATCA GTTCAAGGGA AGTCGGGAGA   3600
CCTGATTTCT AATACTATAT TTTTCCTTTA CAAGCTGAGT AATTCTGAGC AAGTCACAAG   3660
GTAGTAACTG AGGCTGTAAG ATTACTTAGT TTCTCCTTAT TAGGAACTCT TTTTCTCTGT   3720
GGAGTTAGCA GCACAAGGGC AATCCCGTTT CTTTTAACAG GAAGAAAACA TTCCTAAGAG   3780
TAAAGCCAAA CAGATTCAAG CCTAGGTCTT GCTGACTATA TGATTGGTTT TTTGAAAAAT   3840
CATTTCAGCG ATGTTTACTA TCTGATTCAG AAAATGAGAC TAGTACCCTT TGGTCAGCTG   3900
TAAACAAACA CCCAGTTGTA AATGTCTCAA GTTCAGGCTT AACTGCAGAA CCAATCAAAA   3960
AGAATAGAAT CTTTAGAGCA AACTGTGTTT CTCCACATCT GGAGGTGAGT CTGCCAGGGC   4020
AGTTTGGAAA TATTTACTTC ACAAGTATTG ACACTGTTGT TGGTATTAAC AACATAAAGT   4080
TGCTCAAAGG CAATCATTAT TTCAAGTGGC TTAAAGTTAC TTCTGACAGT TTTGGTATAT   4140
TTATTGGCTA TTGCCATTTG CTTTTTGTTT TTTCTCTTTG GGTTTATTAA TGTAAAGCAG   4200
GGATTATTAA CCTACAGTCC AGAAAGCCTG TGAATTTGAA TGAGGAAAAA ATTACATTTT   4260
TGTTTTTACC ACCTTCTAAC TAAATTTAAC ATTTTATTCC ATTGCGAATA GAGCCATAAA   4320
CTCAAAGTGG TAATAACAGT ACCTGTGATT TTGTCATTAC CAATAGAAAT CACAGACATT   4380
TTATACTATA TTACAGTTGT TGCAGATACG TTGTAAGTGA AATATTTATA CTCAAAACTA   4440
CTTTGAAATT AGACCTCCTG CTGGATGTTG TTTTTAACAT ATTAATAAAA CATGTTTAAA   4500
ATTTTGATAT TTTGATAATC ATATTTCATT ATCATTTGTT TCCTTTGTAA TCTATATTTT   4560
ATATATTTGA AAACATCTTT CTGAGAAGAG TTCCCCAGAT TTCACCAATG AGGTTCTTGG   4620
CATGCACACA CACAGAGTAA GAACTGATTT AGAGGCTAAC ATTGACATTG GTGCCTGAGA   4680
TGCAAGACTG AAATTAGAAA GTTCTCCCAA AGATACACAG TTGTTTTAAA GCTAGGGGTG   4740
AGGGGGGAAA TCTGCCGCTT CTATAGGAAT GCTCTCCCTG GAGCCTGGTA GGGTGCTGTC   4800
CTTGTGTTCT GGCTGGCTGT TATTTTTCTC TGTCCCTGCT ACGTCTTAAA GGACTTGTTT   4860
GGATCTCCAG TTCCTAGCAT AGTGCCTGGC ACAGTGCAGG TTCTCAATGA GTTTGCAGAG   4920
TGAATGGAAA TATAAACTAG AAATATATCC TTGTTGAAAT CAGCACACCA GTAGTCCTGG   4980
TGTAAGTGTG TGTACGTGTG TGTGTGTGTG TGTGTGTGTG TGTAAAACCA GGTGGAGATA   5040
TAGGAACTAT TATTGGGGTA TGGGTGCATA AATTGGGATG TTCTTTTTAA AAAGAAACTC   5100
CAAACAGACT TCTGGAAGGT TATTTTCTAA GAATCTTGCT GGCAGCGTGA AGGCAACCCC   5160
CCTGTGCACA GCCCCACCCA GCCTCACGTG GCCACCTCTG TCTTCCCCCA TGAAGGGCTG   5220
GCTCCCCAGT ATATATAAAC CTCTCTGGAG CTCGGGCATG AGCCAGCAAG G
```

Sub a' cont-

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

AACTATTATT GGGGTATGG 23

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

TTGGTGAGGC TTCCTCTGG 19

Sub a' cont.